

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.5
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 13:21:49 ; Search time 1037 Seconds
(without alignments)
5107.725 Million cell updates/sec

Title: US-09-762-105A-14
Perfect score: 182
Sequence: 1 gagctcgcccccggcgtc.....tgacttggacaggctacg 182

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:
2: gb_btg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_uu:
14: gb_vt:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mnu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_uu:
29: em_vt:
30: em_htg_hum:
31: em_htg_in:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rdn:
36: em_htg_mam:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:
Pred. No. 18 the number of results predicted by chance to have a

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
c 1	182	100.0	5270	12	AF176637	Plastid t	AF176637 Plastid t
c 2	111.2	61.1	168	6	AX06661		AX06661 Sequence
c 3	85.8	47.1	127	6	AX137514		AX137514 Sequence
c 4	85.8	47.1	127	6	E51179		E51179 Method for
c 5	85.8	47.1	127	6	E51179		E51179 Process for
c 6	85.8	47.1	767	8	TOBGTGVR		TOBGTGVR
c 7	85.8	47.1	2113	8	CHNTRNI		CHNTRNI
c 8	85.8	47.1	3274	8	CHNPTRNVI		CHNPTRNVI
c 9	85.8	47.1	7829	8	CPY8934		CPY8934
c 10	85.8	47.1	155939	8	CHNTRX		CHNTRX
c 11	85.8	47.1	155939	8	CHNTRX		CHNTRX
c 12	85.8	47.1	156687	8	ABE316582		ABE316582
c 13	85.8	47.1	156687	8	ABE316582		ABE316582
c 14	85.4	46.9	184	6	ARI45914		ARI45914 Sequence
c 15	85.4	46.9	300	6	ARI171710		ARI171710 Sequence
c 16	85.4	46.9	300	6	ARI171711		ARI171711 Sequence
c 17	85.4	46.9	2962	12	XXU12809		XXU12809 Transformat
c 18	85.4	46.9	3019	12	XXU12810		XXU12810 Transformat
c 19	85.4	46.9	3019	12	XXU12811		XXU12811 Transformat
c 20	85.4	46.9	3136	12	AF061065		AF061065 Plastid t
c 21	85.4	46.9	4126	12	XXU12814		XXU12814 Transformat
c 22	85.4	46.9	4126	12	XXU12815		XXU12815 Transformat
c 23	85.4	46.9	4174	12	XXU12812		XXU12812 Transformat
c 24	85.4	46.9	4174	12	XXU12813		XXU12813 Transformat
c 25	85.4	46.9	4304	12	AY005806		AY005806 Transform
c 26	85.4	46.9	7454	12	CTR312391		CTR312391 Chloropla
c 27	85.4	46.9	7626	12	CTR312392		CTR312392 Chloropla
c 28	85.4	46.9	7626	12	CTR312393		CTR312393 Chloropla
c 29	84.8	46.6	159	12	M23201		M23201 Synthetic c
c 30	84.8	46.6	310	12	M23200		M23200 Synthetic c
c 31	84.8	46.6	2731	1	PET31F1P		X62498 E.coli Plas
c 32	84.8	46.6	2829	12	CVPSET6C		CVPSET6C Cloning vec
c 33	84.8	46.6	2829	12	CVPSET6C		CVPSET6C Cloning vec
c 34	84.8	46.6	2830	12	CVPSET5B		CVPSET5B Cloning vec
c 35	84.8	46.6	2830	12	CVPSET6B		CVPSET6B Cloning vec
c 36	84.8	46.6	2831	12	CVPSET5A		CVPSET5A Cloning vec
c 37	84.8	46.6	2831	12	CVPSET6A		CVPSET6A Cloning vec
c 38	84.8	46.6	2829	12	CVT1NDE		CVT1NDE Expression
c 39	84.8	46.6	3984	6	I115353		I115353 Sequence 1
c 40	84.8	46.6	3984	6	X65317		X65317 Cloning vec
c 41	84.8	46.6	3995	12	XPGEMEX2		XPGEMEX2 cloning vec
c 42	84.8	46.6	5231	6	AX001273		AX001273 Sequence
c 43	84.8	46.6	5953	6	AX191649		AX191649 Sequence
c 44	84.8	46.6	39937	7	T7CG		T7CG Genome of b
c 45	83.2	45.7	187	6	AR004778		AR004778 Sequence

ALIGNMENTS

RESULT 1	AF176637/	1					
LOCUS	AF176637		5270 bp	DNA	linear	SYN 24-APR-2000	
DEFINITION			Plastid transformation vector	pMSK49			
ACCESSION	AF176637		Plastid targeting region.	pMSK49			
VERSION	AF176637.1		GI:76737848				
KEYWORDS							
ORGANISM							
ARTIFICIAL SEQUENCES							
REFERENCE	1		(bases 1 to 5270)				
AUTHORS	Khan, M. S. and Maligka, P.						
TITLE	Transformation in higher plants						
JOURNAL	Nat. Biotechnol. 17 (9), 910-915 (1999)						

M protein; PSII N-protein; PSII T-protein; replication origin; ribosomal protein L14; ribosomal protein L16; ribosomal protein L22; ribosomal protein L23; ribosomal protein L24; ribosomal protein L25; ribosomal protein L26; ribosomal protein L31; ribosomal protein S14; ribosomal protein S15; ribosomal protein S16; ribosomal protein S3; ribosomal protein S4; ribosomal protein S5; ribosomal protein S8; ribulose bisphosphate carboxylase; RNA polymerase; small plastid RNA; unidentified reading frame.

ORGANISM	SOURCE	Common tobacco.	Plastid <i>Nicotiana tabacum</i>	<i>Eukaryota</i> ; <i>Viridiplantae</i>	<i>Streptophyta</i> ; <i>Embryophyta</i> ; <i>Tracheophyta</i> ; <i>Liliopsida</i> ; <i>Gramineae</i> .
----------	--------	-----------------	----------------------------------	---	---

REFERENCE AUTHORS TITLE
 Spermatophyta: Magnoliophyta; eudicots; Solanales; Solanaceae; Nicotiana.
 Asterids I; Solanaceae; Nicotiana.
 Asterids I; Solanaceae; Nicotiana.
 Asterids I to 155844
 1 (bases 1 to 155844)
 Shinozaki, K., Ohme, M., Tanaka, M., Wakisugi, T., Hayashida, N.,
 Matsubayashi, T., Zaia, N., Chunwongse, J., Obokata, J.,
 Yamada, H., Shinozaki, K., Ohto, C., Toriyama, K., Meng, B. Y., Sugita, M.,
 Denchi, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A.,
 Tondoh, N., Shimada, H., and Suguri, M.
 The complete nucleotide sequence of tobacco chloroplast genome: its
 gene organization and expression
 1996.

JOURNAL **EMBO J.** 5, 2043-2053 (1986),
2
AUTHORS **Shimada, H. and Sugiyama, M.**
TITLE **Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes of *Arabidopsis thaliana* and *Chlamydomonas reinhardtii***

JOURNAL Nucleic Acids Res. 19 (5), 963-973 (1991)
MEDLINE 91212240
PUBMED 1708498

REFERENCES 3
AUTHORS 3
Mitschelson, K. and Stephen, J.
transfer RNA precursors in tobacco

JOURNAL Nucleic Acids Res. 19 (11), 3150 (1991)
TITLE Processing of histidine triad nucleic acid in chloroplasts

MEDLINE 91279477
PUBMED 2057370
PUBMED-
SERIALS 31640 to 32290; 90230 to 91610)

REFERENCE 4 *Yousef, S.A. and Timmis, J.N.*
AUTHORS *Yalifive, M.A. and Timmis, J.N.*
TITLE *Plastid DNA sequence homologies in the tobacco nuclear genome*

JOURNAL Mol. Gen. Genet. 236 (1): 105-112 (1992)
 MEDLINE 931556674
 DBERMD 1337369

REFERENCE 5 (bases 1 to 155844)
AUTHORS Morton, B.R. and Clegg, M.T.
NOTES *and additional* hotspot and gene conversion in a

TITLE A chloroplast DNA multicopy noncoding region near *rbcL* in the grass family (Poaceae)
JOURNAL *Curr. Genet.* 24 (4), 357-365 (1993)

94073993
MEDLINE
8252646
PUBMED

REFERENCE 6 (1981) 1-20. Lacroix, R. and Lejeune, B.
AUTHORS Nadot, S., Bitar, G., Carter, L.,
TITLE A phylogenetic analysis of monocotyledons based on the chloroplast DNA. A numerical phenetics method

JOURNAL OF MOLECULAR PHYLIOGENETIC EVOLUTION, 4 (3), 257-282 (1995)
0898-1318/95/03257-26\$04.00/0
© 1995 Kluwer Academic Publishers. Printed in Belgium.

PUBMED 8845963
REFERENCE 7 (bases 1 to 155844),
MATERIALS AND METHODS

AUTHORS Kunimimalayaan, M. and Nielsen, B.L.
TITLE Fine mapping of replication origins (ori A and ori B) in Nicotiana tabacum chloroplast DNA

JOURNAL Nucleic Acids Res. 25 (18), 3681-3686 (1997)
MEDLINE 97426512

PUBMED 9278490
REFERENCE 8
AUTHORS Sugiyura, M.

2000-07-27 12:00:00 Direct Submission
2000-07-27 12:00:00 Submitted (12-AUG-1986)
2000-07-27 12:00:00 Accepted (14-AUG-1986)
2000-07-27 12:00:00 JOURNAL (12-AUG-1986)

Submitted (27-FEB-1998), T. Tsuzuki, Data Processing Center, Aichi-Gakuen University, 12 Araiike, Iwasaki, Nisshin, Aichi 470-0195, JAPAN
 On or before Jan 17, 2002 this sequence version replaced gi:1264799
 gi:18807.
 The circular tobacco chloroplast DNA sequence is presented in a linearized form by cutting at the junction (J1A) between IRA and LSC.
 J1A is designated zero and numbered proceeding towards LSC. The DNA strand which codes for the large subunit of ribulose 1,5-bisphosphate carboxylase is designated as A strand and the complementary strand as B strand.
 Nucleotide sequence of the B strand is presented.
 Large single copy region (LSC): 1 - 86686 (86686 bp) Invert repeat B (IRB) : 86687 - 112027 (2541 bp) Small single copy region (SSC): 112028 : 86687 - 112027 (2541 bp) Invert repeat A (IRA) : 112028 - 130598 (1851 bp) Invert repeat A (IRA) : 130599 - 155939 (2531 bp) An alphabetical index of tobacco chloroplast genes and ORFs GENE NUCLEOTIDE NUMBER (FROM accb 5973 112862
 ars1 14569
 ars2 12148
 atpa 56777
 atpb 55284
 atpe 13452
 atpf 14099
 atpi 16000
 c1pp 74507
 c1pp 82495
 infa 3658
 matK 129293
 ndha 992621,
 ndhc 52667
 ndhb 119025
 ndhd 111955
 ndhe 111292
 ndhf 120109
 ndhg 125116
 ndhh 121609
 ndhi 51465
 ndhj 52425
 ndhk 48341
 ORF70A 102102,
 ORF70B 46412
 ORF74 110824,
 ORF75 131802
 ORF79 96536,
 ORF92 96119,
 ORF99 66176
 ORF103 67588
 ORF105 37566
 ORF115 96247
 ORF131 102346,
 ORF350 111029
 oria 104769,
 orib 111778,
 petA 64335
 petB 77452
 petD 79043
 petG 68570
 petL 68293
 psaa 43486
 psab 41208
 psac 119389
 psai 62083
 psaj 69565
 psba 1595
 psbb 74953
 psbc 35515
 psbd 34470
 psbe 67729
 psbf 66848
 psbh 77101
 psbi 8198

TITLE

psbJ	66485	L23; ribosomal protein L33; ribosomal protein L33; ribosomal protein L36; ribosomal protein S11; ribosomal protein S12; ribosomal protein S14; ribosomal protein S15; ribosomal protein S16; ribosomal protein S18; ribosomal protein S19; ribosomal protein S2; ribosomal protein S3; ribosomal protein S4; ribosomal protein S7; ribosomal protein S8; ribulose bisphosphate carboxylase; RNA Polymerase; small plastid RNA; unidentified reading frame.
psbK	7835	
psbL	66726	
psbL	30861	
psbM	76989	
psbN	76680	
psbt	57595	
rbCL	88233	154393
rp112	83546	
rp114	83595	
rp116	71409	
rp120	86350	
rp122	88533	154093
rp123	115661	
rp132	70131	
rp133	82165	
rp136	81468	
rpoA	27111	
rpoB	24293	
rpoC1	21342	
rpoC2	16937	
rps2	83983	
rps3	48141	
rps4	100007	142619
rps7	83006	
rps8	81950	
rps11	100854	141772
3'rp12	572334	
rps14	318881	
rps15	125491	
rps16	6211	
rps18	70518	
rps19	86682	
rrn4.5	10940	1333086
rrn5	10599	133027
rrn16	102761	139885
rrn23	106329	136297
spRA	115294	
trnA-UGC	105394	137732
trnC-GCA	28793	
trnD-GUC	32007	
Query Match	47.1%	Score 85.8; DB 8; Length 155939;
Best Local Similarity	97.8%	Pred. No. 1.0e-16;
Matches	87; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
Qy	5	TCCCTCCCGCGTCGTCGTCATGAGATGGATAAGAGGCTCGGGATTGACGTGAGGG 64
Db	102562	TTGCTCCCCCGCGCTGTCATGAGATGGATAAGAGGCTCGGGATTGACGTGAGGG 102621
Qy	65	GCGAGGGATGGCTATATTCTGGAGGCA 93
Db	102622	GGCAGGGATGGCTATATTCTGGAGGCA 102650
RESULT	11	
CHNTXX/C	CHNTXX	155939 bp DNA circular PLN 27-JUL-2001
LOCUS	Nicotiana tabacum chloroplast genome DNA.	
DEFINITION	Nicotiana tabacum chloroplast genome DNA.	
ACCESSION	200044 554304	
VERSION	200044.1 GI:2924257	
KEYWORDS	16S ribosomal RNA; 23S ribosomal RNA; 4.5S ribosomal RNA; 5S ribosomal RNA; acetyl-CoA carboxylase; ATP-dependent protease; autonomous b/f complex; chloroplast; circular; complete genome; cytochrome b/f complex; cytochrome b6; cytochrome f; initiation factor; inverted repeat; matruse; NADH dehydrogenase; PSI 9kd protein; PSI J-protein; PSI P700 apoprotein; PSI 10kd phosphoprotein; PSI 32kd protein; PSI 44kd protein; PSI 47kd protein; PSI D2 protein; PSI L-protein; PSI M-protein; PSI J-protein; PSI T-protein; replication origin; ribosomal protein L14; ribosomal protein L16; ribosomal protein L22; ribosomal protein L20; ribosomal protein L22; ribosomal protein	

SOURCE	ORGANISM	Plastid Nicotiana tabacum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.	
REFERENCE	AUTHORS	Shinozaki, K., Ome, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaia, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takiwa, F., Kato, A., Tohdo, N., Shimada, H., and Sugiura, M.
TITLE		The complete nucleotide sequence of tobacco chloroplast genome: its gene organization and expression
JOURNAL	EMBO J. 5, 2043-2049 (1986)	
REFERENCE	AUTHORS	Shimada, H. and Sugiura, M.
	TITLE	Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes
JOURNAL	NUCLIC ACIDS RES. 19 (5), 983-995 (1991)	
REFERENCE	AUTHORS	Mitchelson, R. and Stephen, J.
	TITLE	Processing of histidine t-transfer RNA precursors in tobacco chloroplasts
JOURNAL	NUCLIC ACIDS RES. 19 (11), 3150 (1991)	
REFERENCE	AUTHORS	Shimada, H. and Sugiura, M.
	TITLE	Fine structural features of the chloroplast genome: comparison of the sequenced chloroplast genomes
JOURNAL	NUCLIC ACIDS RES. 19 (5), 983-995 (1991)	
REFERENCE	AUTHORS	Ayliffe, M. A. and Timmis, J. N.
	TITLE	Plastid DNA sequence homologies in the tobacco nuclear genome noncoding region near rbcL in the grass family (Poaceae)
JOURNAL	CURR. GENET. 24 (4), 357-365 (1993)	
REFERENCE	AUTHORS	Nadot, S., Bittar, G., Carter, J., Lacroix, R. and Lejeune, B.
	TITLE	A phylogenetic analysis of monocotyledons based on the chloroplast gene rps4, using parsimony and a new numerical phenetics method
JOURNAL	MOL. GEN. GENET. 236 (1), 105-112 (1992)	
PUBLMED	94073993	
REFERENCE	AUTHORS	Nadot, S., Bittar, G., Carter, J., Lacroix, R. and Lejeune, B.
	TITLE	A phylogenetic analysis of monocotyledons based on the chloroplast gene rps4, using parsimony and a new numerical phenetics method
JOURNAL	MOL. PHYLOGENET. EVOL. 4 (3), 257-282 (1995)	
PUBLMED	8845963	
REFERENCE	AUTHORS	Runnimalaiyan, M. and Nielsen, B. L.
	TITLE	Fine mapping of replication origins (ori A and ori B) in Nicotiana tabacum chloroplast DNA
JOURNAL	NUCLIC ACIDS RES. 25 (18), 3681-3686 (1997)	
PUBLMED	9726512	
REFERENCE	AUTHORS	Sugiura, M.
	TITLE	Direct Submission
JOURNAL	NUCLIC ACIDS RES. 9 (bases 1 to 155939)	
PUBLMED	9278490	
REFERENCE	AUTHORS	Runnimalaiyan, M. and Nielsen, B. L.
	TITLE	Fine mapping of replication origins (ori A and ori B) in Nicotiana tabacum chloroplast DNA
JOURNAL	NUCLIC ACIDS RES. 25 (18), 3681-3686 (1997)	
PUBLMED	9726512	
REFERENCE	AUTHORS	Sugiura, M.
	TITLE	Direct Submission
JOURNAL	NUCLIC ACIDS RES. 9 (bases 1 to 155939)	
PUBLMED	9278490	
REFERENCE	AUTHORS	Tsuduki, T.
	TITLE	Direct Submission
JOURNAL	NUCLIC ACIDS RES. 9 (bases 1 to 155939)	
PUBLMED	9726512	
REFERENCE	AUTHORS	Aichi Gakuin University, 12 Araike, Iwasaki, Nisshin, Aichi 470-0195, JAPAN

COMMENT
On or before Jan 17, 2002 this sequence version replaced gi:264799.
gi:11807.
The circular tobacco chloroplast DNA sequence is presented in a
linearized form by cutting at the junction (JLA) between IRA and
LSC.
JLA is designated zero and numbered proceeding towards LSC. The DNA
strand which codes for the large subunit of ribulose
1,5-bisphosphate carboxylase is designated as A strand and the
complementary strand as B strand.
Nucleotide sequence of the B strand is presented.
Large single copy region (LSC): 1 - 86686 (86686 bp) Inverted
repeat B (IRB): 86687 - 112027 (25341 bp) Small single copy
region (SSC): 112028 - 130598 (18571 bp) Inverted repeat A (IRA)
: 130599 - 155939 (25411 bp) An alphabetical index of tobacco
chloroplast genes and ORFs GENE NUCLEOTIDE NUMBER (FROM
accD 59193
accP 112662
ars1 14369
ars2 12448
atpA 56777
atpB 55784
atpB 13452
atpF 14099
atpH 16000
atpI 74507
clpp 82495
infA 36558
mack 123933
ndhA 99261, 1433355
ndhB 52667
ndhC 119025
ndhD 119955
ndhE 114292
ndhF 120709
ndhG 125116
ndhH 121609
ndhI 51465
ndhJ 52425
ORF77A 48941
ORF77B 102102,
ORF77 46472,
ORF75 110824,
ORF79 96556,
ORF92 96119,
ORF99 66176
ORF103 67588
ORF105 37566
ORF115 96470,
ORF131 103446,
ORF350 111029,
oria 10769,
orib 111778,
petA 63335
petB 77452
petC 79043
petD 68370
petG 68223
petL 43486
psAA 41108
psAB 11989
psAC 62083
psAI 69585
psAJ 1595
psB 74953
psB 35515
psCD 34470
psD 67129
psEB 66868
psBH 77101
psB 8398
psB 66485
psB 7835
psB 66735
psB

RESULT 12											
LOCUS	ABE316582	156687 bp	DNA	circular PLN 05-APR-2002							
DEFINITION	Atropa belladonna Complete Plastid Chromosome, strain Ab5p (kan).										
VERSION	AJ316582.1	GI:20068310	4.5S ribosomal RNA; 4.5S ribosomal RNA; 5S								
KEYWORDS	16S ribosomal RNA; accD gene; acetyl-CoA carboxylase beta subunit; ATPase ribosomal RNA; atpC gene; catalytic subunit; ccsA gene; ccsB protein ATP-dependent protease; atpA gene; ATPase alpha subunit; ATPase beta subunit; ATPase epsilon subunit; ATPase subunit I; ATPase subunit III; ATPase subunit IV; ttpB gene; atpE gene; atpF gene; atpH gene; atpI gene; atpJ gene; catalytic subunit; ccsA gene; ccsB protein required for heme attachment to c-type cytochrome genes; ccmA gene; cllP gene; Cyt b6/f complex subunit IV; cyt b6/f complex subunit VI; cytochrome b/f complex subunit V; cytochrome b6; cytochrome b6/f complex subunit IV; cytochrome f; matK gene; matK gene; NADH dehydrogenase ND6 subunit; NADH dehydrogenase 18kD subunit; NADH dehydrogenase 19kD subunit; NADH dehydrogenase 32kD subunit; NADH dehydrogenase 49kD subunit; NADH dehydrogenase ND1 subunit; NADH dehydrogenase ND2 subunit; NADH dehydrogenase ND4 subunit; NADH										
Query Match	47.1%	Score 85.8;	DB 8;	Length 155939;							
Best Local Similarity	97.8%	Pred. No. 1	6e-11;								
Matches	87;	Conservative	0;	Matches	1;	Indels	0;	Gaps			
/ 5 TGCCTCCCGCGCGTCGTCATGAGATGGATAAGGCCCTGTCGGATGACGTAGGG 64											
/ 65 GCGAGGGATGGCTATATTTCTGGAGGGA 93											
/ 14 00044 TGCCTCCCGCGCGTCGTCATGAGATGGATAAGGCCCTGTCGGATGACGTAGGG 139976											
/ 14 00044 GCGAGGGATGGCTATATTTCTGGAGGGA 139976											

ORGANISM	SOURCE	PLANT
<i>belladonna</i>		
<i>Atropa belladonna</i>		
<i>Eukarya</i>		
<i>Viridiplantae</i>		
<i>Streptophyta</i>		
<i>Embryophyta</i>		
<i>Tracheophyta</i>		
<i>Spermatophyta</i>		
<i>Magnoliophyta</i>		
<i>eudicotyledons</i>		
<i>core eudicots</i>		

Asteridae: euasterids I; Solanales; Solanaceae; Atropa.
1 Schmitz-Linneweber, C., Regel, R., Gia Du, T., Hupfer, H.,
Herrmann, R.G. and Maier, R.M.
The nucleotide sequence of the plastid chromosome of *Atropa*
belladonna (deadly nightshade) and its comparison with that of
Nicotiana tabacum with emphasis on sequence elements relevant for

microevolution
Unpublished
2 (bases 1 to 156687)
Schmitz-Linneweber, C.
Direct Submission
Submitted (24-MAY-2000) Schmitz-Linneweber C., Botanisches
Institut, Ludwig Maximilians Universitaet Muenchen, Menzinger Str.
67, Muenchen, Bavaria 80638, Germany
Location: <http://www.boku.ac.at/botanik/linneweber/>

CATALOGUE

source			
	CDS		
		gene	
			CDS
			gene
			tRNA
			gene
			CDS
			gene
			tRNA

exon intron gene CDS

```

1. 156687
/organism="Atropa belladonna"
/strain="Ab5p(kan)"
/db_xref="Taxon:33113"
complement(join(100334. .1000895. .101126,
/codon_start=1
/product="ribosomal protein S12"
/protein_id="CAC88091.1"
/db_xref="GI:20068378"
/translation="MPTINOLIRNTRPQIRNVTKSPALRGCPQRGTCTRVYITTPKK
PNSALKVAVRPTISCEFEITAYIPIGHNLQEHSSVVLVGGRVKDPGLGVRHVGTL
DAVGKDRQGRSKRSGVKKP
join_complement(100895. .101126),
complement(100360. .100894),complement(100333. .100359),
complement(72478. .72591),.142431. .142662,.143198. .143223,
/gene="rps12"
join_complement(72478. .72591),.142431. .142662,
143198. .143223)
/gene="rps12"
/codon_start=1
/product="ribosomal protein S12"
/protein_id="CAC88068.1"
/db_xref="GI:20068355"
/translation="MPTINOLIRNTRPQIRNVTKSPALRGCPQRGTCTRVYITTPKK
PNSALKVAVRPTISCEFEITAYIPIGHNLQEHSSVVLVGGRVKDPGLGVRHVGTL
DAVGKDRQGRSKRSGVKKP
join_complement(5. .82),
/gene="tRNA-his (GUG)"
/gene="tRNA-his (GUG)"
/product="tRNA-His"
/note="codon recognized: CAC"
complement(494. .1555)
/gene="psba"
complement(494. .1555)
/gene="psba"
/codon_start=1
/product="PSII 32 kd protein"
/protein_id="CAC88024.1"
/db_xref="GI:20068311"
/translation="MTAIERRERESSLWGRFCNWTISTBNRLYCGWFGVIMIPTLLTA
TSVFLIAFIAPPDIDGHEPVSSOLLGNNIISGAIIPTSAAGLHYPINEAASY
DEWLWYDGFVPLVHLGLGVGATVYDQEMTIVQAEINILMHPFLMGLVAGVAYSAVTAATAVFLI
YPIGQFSDFDMPGLGSGTENMIVQAEINILMHPFLMGLVAGVAYSAVTAATAVFLI
AVSPLVNGTENESANGYRFGQEEETYVIAAGYFGGLFOASFNNSRSUHFELA
AWPFLVNGTENESANGYRFGQEEETYVIAAGYFGGLFOASFNNSRSUHFELA
NAHNFPLDAAIEPTNG"
complement(join(1769. .1804. .1805. .4323. .4324. .4360))
/gene="tRNA-Lys (UUU)"
complement(join(1769. .1804. .4324. .4360))
/product="tRNA-Lys (UUU)"

```

/Product: tRNA-Lys
/note= "codon recognized: AAA"
complement(1769 . 1804)
/gene="tRNA-Lys (UUU)"
/number=2
complement(1805 . 4323)
/gene="tRNA-Lys (UUU)"
complement(2068 . 3597)
/gene="matK"
complement(2068 . 3597)
/gene="matK"
/codon_start=1

```

<product> "maturase"
<protein_id> CAC8025.1
<db_xref> "GT1:20068312"
<translation> "MEETIQRQLQPDSSQHNFYLIFQEQTYIYALADHGGLNINSIL
LVEPGINNNSTLIVKRLITTYOONHFLISTDNDKNTPLGCKNSLYSMISGDFAF
IPEVPLRSLSSSEGGKLLKISNRHSTPLKEMNNSHANVYDILIPYF
LEIYOTLRYWKDASSLHLRRFFGEHYNWNSLTSITKPGYSEFKKNRFFEFLYNS
YVECSTFEVLRNQSHLAPTSFGALLERYFYGKLERLVEFADQYTLWLEKDPL

```

REFERENCE	JOURNAL	ORGANISM	FEATURES
Schmitz-Linnweber, C., Regel, R., Gia Du, T., Hupfer, H., Hermann, R.G. and Maier, R.M.	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Asteridae; eudasterids I; Solanales; Solanaceae; Atropa.		
1. Schmitz-Linnweber, C. Direct Submission	The nucleotide sequence of the plastid chromosome of <i>Atropa belladonna</i> (deadly nightshade) and its comparison with that of <i>Nicotiana tabacum</i> with emphasis on sequence elements relevant for microevolution		
2. (bases 1 to 156687)	Unpublished		
Schmitz-Linnweber, C. Submitter	Ludwig-Maximilians Universitaet Muenchen, Muenchen, Bavaria 80638, Germany		
1. 156687	Local/Qualifies		
Strain="Atropa belladonna"	Strain="Abp(kan)"		

Db 2 GCTCCCCGGCGTCAATGAAATGGATAAGGGCTCCGGGATGTGAGGGGG 61
 Qy 67 CAGGGATGGCTATATTCTGGAGGG 93
 ||||||| ||||| ||||| ||||| |||||
 Db 62 CAGGGATGGCTATATTCTGGAGCGA 88

RESULT 15
 ARI17170 AR171710 Sequence 4 from patent US 6297054. DNA linear PAT 17-DEC-2001
 LOCUS DEFINITION Sequence 4 from patent US 6297054.
 ACCESSION AR171710
 VERSION AR171710.1 GI:17910660
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.
 1 (bases 1 to 300)

REFERENCE Maliga, P., Carrer, H. and Chaudhuri, S.
 AUTHORS Maliga, P., Carrer, H. and Chaudhuri, S.
 TITLE Editing-based selectable plastid marker genes
 Patent: US 6297054-A 4 02-OCT-2001;
 FEATURES Location/Qualifiers
 1..300
 SOURCE /organism="unknown"

BASE COUNT 90 a 63 c 86 g 61 t
 ORIGIN

Query Match 46.9%; Score 85.4; DB 6; Length 300;
 Best Local Similarity 98.9%; Pred. No. 1..1e-16;
 Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GCTCCCCGGCGTCAATGAAATGGATAAGGGCTCCGGGATGTGAGGGGG 66
 Db 17 GCTCCCCGGCGTCAATGAAATGGATAAGGGCTCCGGGATGTGAGGGGG 76

Qy 67 CAGGGATGGCTATATTCTGGAGGG 93
 ||||||| ||||| ||||| |||||

Db 77 CAGGGATGGCTATATTCTGGAGCGA 103

Search completed: May 25, 2003, 14:05:05
 Job time : 1177 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 13:38:09 ; Search time 61 Seconds
(without alignments)
915.003 Million cell updates/sec

Title: US-09-762-105A-14
Perfect score: 182
Sequence: 1 gagctcgctccccccggcgtc.....tgactgtggacaggcttagc 182

Scoring table: IDENTITY_NUC
GapOp 10.0 , GapExt 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:
1: /cg2_6/ptodata/2/ina/5A_COMBO.seq;*
2: /cg2_6/ptodata/2/ina/5B_COMBO.seq;*
3: /cg2_6/ptodata/2/ina/6A_COMBO.seq;*
4: /cg2_6/ptodata/2/ina/6B_COMBO.seq;*
5: /cg2_6/ptodata/2/ina/BCTUS_COMBO.seq;*
6: /cg2_6/ptodata/2/ina/backfiles1.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	89.6	49.2	171	1 US-09-217-360-16	Sequence 16, Appl
2	86.4	47.5	129	2 US-08-189-256A-25	Sequence 25, Appl
3	86.4	47.5	129	4 US-09-193-853-25	Sequence 25, Appl
4	86	47.3	140	2 US-08-189-256A-19	Sequence 19, Appl
5	86	47.3	140	4 US-09-193-853-19	Sequence 19, Appl
6	86	47.3	164	2 US-08-189-256A-26	Sequence 26, Appl
7	86	47.3	164	4 US-09-193-853-26	Sequence 26, Appl
8	85.4	46.9	161	2 US-08-189-256A-18	Sequence 18, Appl
9	85.4	46.9	161	4 US-09-193-853-18	Sequence 18, Appl
10	85.4	46.9	165	2 US-08-189-256A-4	Sequence 4, Appl
11.	85.4	46.9	165	4 US-09-193-853-4	Sequence 4, Appl
12.	85.4	46.9	168	2 US-08-189-256A-2	Sequence 2, Appl
13	85.4	46.9	168	4 US-09-193-853-2	Sequence 2, Appl
14	85.4	46.9	184	4 US-09-283-419-3	Sequence 3, Appl
15	85.4	46.9	258	2 US-08-193-256A-24	Sequence 24, Appl
16	85.4	46.9	258	4 US-09-193-853-24	Sequence 24, Appl
17	85.4	46.9	300	4 US-09-202-316-4	Sequence 4, Appl
18	85.4	46.9	300	4 US-09-202-316-7	Sequence 7, Appl
c 19	85.4	46.9	1134	2 US-08-189-256A-10	Sequence 10, Appl
c 20	85.4	46.9	1134	4 US-09-193-853-10	Sequence 10, Appl
c 21	85.4	46.9	1143	4 US-09-142-114B-6	Sequence 6, Appl
c 22	85.4	46.9	1208	2 US-08-189-256A-28	Sequence 28, Appl
c 23	85.4	46.9	1208	4 US-09-193-853-28	Sequence 28, Appl
c 24	85.4	46.9	1416	2 US-08-189-256A-27	Sequence 27, Appl
c 25	85.4	46.9	1416	4 US-09-193-853-27	Sequence 27, Appl
c 26	85.4	46.9	1417	4 US-09-142-114B-7	Sequence 7, Appl
c 27	85.4	46.9	2962	2 US-08-189-256A-3	Sequence 3, Appl

RESULT 1
US-08-217-360-16
; Sequence 16, Application US/08217360
; Patent No. 553091
; GENERAL INFORMATION:
; APPLICANT: MALIGA, Pal
; TITLE OF INVENTION: METHODS FOR PRODUCING CYTOPLASMIC
; MALE STERILITY IN PLANTS AND USE THEREOF IN PRODUCTION OF H
; TITLE OF INVENTION: SEED
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Danno, Dorfman, Herrrell and Skillman,
; ADDRESS: P.C.
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217-360
; CLASSIFICATION: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, Jane E.
; REGISTRATION NUMBER: 36, 252
; REFERENCE/DOCKET NUMBER: Rutgers University
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-217-360-16
; Query Match 49.2%; Score 89.6%; DB 1; Length 171;
; Best Local Similarity 78.3%; Pred. No. 1.4e-1;
; Matches 130; Conservative 0; Mismatches 14; Indels 22; Gaps 1;
; Qy 7 GCTCCGCCGCCGTCATGAGATGGATAAGAGCTGAGGGGG 1;

Db	1	GCTCCCGCCGCTGTCATAATGAAATGGATAAGGGCTCGGGATTTGAGTCGAGGGG	60
Qy	67	CAGGGATGGCTATATTCCTG-	104
Db	61	CAGGGATGGCTATATTCCTGAAATTGGAGGGATACAAATTGGAGCGAATCTGGCGCA	120
Qy	105	TTCGCCACTAGAAATAATTTCGTTAACTTAAAGACCGAGATAAC	150
Db	121	ATTAGATCTGAAATAATTGTTAACTTAAAGAGATAAC	166

```

RESULT 3
US-09-193-853-25
US Sequence 25, Application US/09193853
    Sequence 25, Application US/09193853
    Patent No. 6,388168
    GENERAL INFORMATION:
    APPLICANT: Maliga, Pal
    APPLICANT: Svab, Zora
    APPLICANT: Staub, Jeffrey
    APPLICANT: Zoubenko, Oleg V.
    APPLICANT: Allison, Lori A.
    APPLICANT: Carrer, Helaine
    APPLICANT: Kanyevski, Ivan
    TITLE OF INVENTION: DNA Constructs and Methods for Stably
    TITLE OF INVENTION: Transforming Plastids of Multicellular Plants a
    TITLE OF INVENTION: Expressing Recombinant Proteins Therein
    NUMBER OF SEQUENCES: 47
    NUMBER OF SEQUENCES: 47
    CORRESPONDENCE ADDRESS:
    ADDRESSEE: Dainin Dorfman, Herrell and Skillman
    STREET: 1601 Market Street Suite 720
    CITY: Philadelphia
    STATE: PA
    COUNTRY: USA
    ZIP: 19103-2307
    COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/193, 853
    FILING DATE:
    CLASSIFICATION:
    PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 08/189, 256
    FILING DATE:
    PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/518, 763
    FILING DATE: 01-MAY-1990
    ATTORNEY/AGENT INFORMATION:
    NAME: Reed, Janet E.
    REGISTRATION NUMBER: 36, 252
    TELECOMMUNICATION INFORMATION:
    TELEPHONE: (215) 563-4100
    TELEFAX: (215) 563-4044
    INFORMATION FOR SEQ ID NO: 25:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 129 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    US-09-193-853-25

Query Match 47-58: Score 86.4; DB 4; Length 129;
    Best Local Similarity 89.4%; Pred. No. 1 5e-20; Gap 0;
    Matches 93; Conservative 0; Mismatches 11; Indels 0;
    7 GCTCCCCCGCCGCTGTCATGAAATGATAAAGAGGCTCGTGGATGACGTTAGGGGG
    6 GTCGCCCGCCGCTGTCATGAAATGATAAAGAGGCTCGTGGATGACGTTAGGGGG
    67 CAGGGATGCGCTATTTCTGGAGACCAACGGCTCGTGGATGACGTTAGGGGG
    66 CAGGGATGCGCTATTTCTGGAGACCAACTCGGGCAATTCTGGACGAACTCGGGCAATTCTGGAC 110
    66 CAGGGATGCGCTATTTCTGGACGAACTCGGGCAATTCTGGAC 109

```

RESULT 4

US-08-189-256A-19

Sequence 19, Application US/08189256A

Patent No. 5877402

GENERAL INFORMATION:

APPLICANT: Maliga, Pal

APPLICANT: Svab, Zora

APPLICANT: Staub, Jeffrey

APPLICANT: Zoubenko, Oleg V.

APPLICANT: Allison, Lori A.

APPLICANT: Carrer, Helaine

APPLICANT: Kanevskii, Ivan

APPLICANT: Staub, Jeffrey

APPLICANT: Zoubenko, Oleg V.

APPLICANT: Allison, Lori A.

APPLICANT: Carrer, Helaine

APPLICANT: Kanevskii, Ivan

TITLE OF INVENTION: DNA Constructs and Methods for Stably Transforming Plastids of Multicellular Plants and Expressing Recombinant Proteins Therein

TITLE OF INVENTION: DNA Constructs and Methods for Stably Transforming Plastids of Multicellular Plants and Expressing Recombinant Proteins Therein

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dann, Dorfman, Herrell and Skillman

STREET: 1601 Market Street Suite 720

CITY: Philadelphia

STATE: PA

COUNTRY: USA

Z.I.P.: 19103-2307

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/193,853

FILING DATE: 31-JAN-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/111,398

FILING DATE: 25-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/518,763

FILING DATE: 01-MAY-1990

ATTORNEY/AGENT INFORMATION:

NAME: Reed, Janet E.

REGISTRATION NUMBER: 36,252

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 563-4100

TELEFAX: (215) 563-4044

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 140 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-09-193-853-19

Query Match 47.3%; Score 86; DB 2; Length 140;

Best Local Similarity 90.2%; Pred. No. 2.1e-20;

Matches 92; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Db 17 GCTCCCGCGGCTTCAATGGATGATAAGGGCTGTGGATTCAGCTGAGGGGG 118

Qy 7 GCTCCCGCGGCTTCAATGGATGATAAGGGCTGTGGATTCAGCTGAGGGGG 66

Db 17 GCTCCCGCGGCTTCAATGGATGATAAGGGCTGTGGATTCAGCTGAGGGGG 76

Qy 67 CAGGGATGGCTATATTCTGGAGGAGACACACAGGTTC 108

Db 17 GCTCCCGCGGCTTCAATGGATGATAAGGGCTGTGGATTCAGCTGAGGGGG 76

Qy 67 CAGGGATGGCTATATTCTGGAGGAACTCGGGCGAATTC 118

Db 17 GCTCCCGCGGCTTCAATGGATGATAAGGGCTGTGGATTCAGCTGAGGGGG 118

RESULT 5

US-08-189-256A-26

Sequence 26, Application US/08189256A

Patent No. 5877402

GENERAL INFORMATION:

APPLICANT: Maliga, Pal

APPLICANT: Svab, Zora

APPLICANT: Staub, Jeffrey

TITLE OF INVENTION: Transforming Plasmids of Multicellular Plants and Expressing Recombinant Proteins Therein
 TITLE OF INVENTION: Expressing Recombinant Proteins Therein
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dann Dorfman, Herrrell and Skillman
 STREET: 1601 Market Street Suite 720
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103-2307
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/193,853
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/189,256
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/518,763
 FILING DATE: 01-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Janet E.
 REGISTRATION NUMBER: 36,252
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 563-4100
 TELEFAX: (215) 563-4044
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 164 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-09-193-853-26

Query Match Score 47.3%; DB 4; Length 164;
 Best Local Similarity 90.2%; Pred. No. 2.e-20; Indels 0; Gaps 0; Gaps
 Matches 92; Conservative 0; Mismatches 10;

Match	Sequence	Score	Length
Qy	7 GCTCCCGCCGTCATGAGATGATGAGCTCGTGGGATGAGCTCGTGGGAGCTCGTGGGAGGG 66	47	164
Db	6 GCTCCCGCCGTCATGAGATGATGAGCTCGTGGGATGAGCTCGTGGGAGCTCGTGGGAGGG 65	47	163
Qy	67 CAGGGATGGCTATATTTCATGGAGACCAACAGCTTC 108	47	164
Db	66 CAGGGATGGCTATATTTCATGGAGACCAACAGCTTC 107	47	163

RESULT 8
 US-08-189-256A-18
 Sequence 18, Application US/08189256A
 Patent No. 5,877402
 GENERAL INFORMATION:
 APPLICANT: Maliga, Pal
 APPLICANT: Svab, Zora
 APPLICANT: Staub, Jeffrey
 APPLICANT: Zoubenek, Oleg V.
 APPLICANT: Allison, Lori A.
 APPLICANT: Carrer, Helaine
 APPLICANT: Kaneko, Ivan
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dann Dorfman, Herrrell and Skillman

TITLE OF INVENTION: DNA Constructs and Methods for Stably Transforming Plasmids of Multicellular Plants and Expressing Recombinant Proteins Therein
 TITLE OF INVENTION: Expressing Recombinant Proteins Therein
 NUMBER OF SEQUENCES: 47

STREET: 1601 Market Street Suite 720
 CITY: Philadelphia
 STATE: PA
 ZIP: 19103-2307
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 APPLICATION NUMBER: US/08/189,256A
 FILING DATE: 31-JAN-1994
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/111,398
 FILING DATE: 25-AUG-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/518,763
 FILING DATE: 01-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Janet E.
 REGISTRATION NUMBER: 36,252
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 563-4100
 TELEFAX: (215) 563-4044
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 161 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI SENSE: NO
 US-08-189-256A-18

Query Match 46.9%; Score 85.4%; DB 2; Length 161;
 Best Local Similarity 98.9%; Pred. No. 3.6e-20;
 Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 7 GCTCCCCCGCGTCGTCAATGAGATGATAAGAGGCTGTTGGATTCAGTGGATGGGG 66
 Db 23 GCTCCCCCGCGTCGTCAATGAGATGATAAGAGGCTGTTGGATTCAGTGGATGGGG 82
 Qy 67 CAGGGATGGTATATTCGGAGGA 93
 Db 83 CAGGGATGGTATATTCGGAGCGGA 109

RESULT 9
 US-09-193-853-18
 Sequence 18, Application US/09193853
 Patent No. 6388168
 GENERAL INFORMATION:
 APPLICANT: Maliga, Pal
 APPLICANT: Svab, Zora
 APPLICANT: Staub, Jeffrey
 APPLICANT: Zoubenko, Oleg V.
 APPLICANT: Allison, Lori A.
 APPLICANT: Carrer, Helaine
 APPLICANT: Kanevski, Ivan
 TITLE OF INVENTION: DNA Constructs and Methods for Stably
 Transforming Plastids of Multicellular Plants and
 Expressing Recombinant Proteins Therein
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dann, Dorfman, Hurrell and Skillman
 STREET: 1601 Market Street Suite 720
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103-2307

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/193,853
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/189,256
 FILING DATE:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/518,763
 FILING DATE: 01-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Janet E.
 REGISTRATION NUMBER: 36,252
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 563-4100
 TELEFAX: (215) 563-4044
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 161 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI SENSE: NO
 US-09-193-853-18

Query Match 46.9%; Score 85.4%; DB 4; Length 161;
 Best Local Similarity 98.9%; Pred. No. 3.6e-20;
 Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 7 GCTCCCCCGCGTCGTCAATGAGATGATAAGAGGCTGTTGGATTCAGTGGATGGGG 66
 Db 23 GCTCCCCCGCGTCGTCAATGAGATGATAAGAGGCTGTTGGATTCAGTGGATGGGG 82
 Qy 67 CAGGGATGGTATATTCGGAGGA 93
 Db 83 CAGGGATGGTATATTCGGAGCGGA 109

RESULT 10
 US-08-189-256A-4
 Sequence 4, Application US/08189256A
 Patent No. 587402
 GENERAL INFORMATION:
 APPLICANT: Maliga, Pal
 APPLICANT: Svab, Zora
 APPLICANT: Staub, Jeffrey
 APPLICANT: Zoubenko, Oleg V.
 APPLICANT: Allison, Lori A.
 APPLICANT: Carrer, Helaine
 APPLICANT: Kanevski, Ivan
 TITLE OF INVENTION: DNA Constructs and Methods for Stably
 Transforming Plastids of Multicellular Plants and
 Expressing Recombinant Proteins Therein
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dann, Dorfman, Hurrell and Skillman
 STREET: 1601 Market Street Suite 720
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103-2307

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/189,256A
 FILING DATE: 31-JAN-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/111,398
 FILING DATE: 25-AUG-1993
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/518,763
 FILING DATE: 01-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Janet E.
 REGISTRATION NUMBER: 36,252
 TELEPHONE: (215) 563-4100
 TELEFAX: (215) 563-4044
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 165 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-09-193-853-4

Query Match 46.9%; Score 85.4%; DB 4; Length 165;
 Best Local Similarity 98.9%; Pred. No. 3.7e-20; Indels 0; Gaps 0;
 US-08-189-256A-4

Query Match 46.9%; Score 85.4%; DB 2; Length 165;
 Best Local Similarity 98.9%; Pred. No. 3.7e-20; Indels 0; Gaps 0;
 Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy - 7 GCTCCCCGCCCTCGTCAATGAGATGGATAAGGGCTGTTGGATGAGCTGGGG 66
 Db 1 GCTCCCCGCCCTCGTCAATGAGATGGATAAGGGCTGTTGGATGAGCTGGGG 60

RESULT 12
 US-08-189-256A-2
 Sequence 2, Application US/08189256A
 Patent No. 5877402
 GENERAL INFORMATION:
 APPLICANT: Maliga, Pal
 APPLICANT: Svab, Zora
 APPLICANT: Staub, Jeffrey
 APPLICANT: Zoubenko, Oleg V.
 APPLICANT: Allison, Lori A.
 APPLICANT: Kanevski, Ivan
 APPLICANT: Carrier, Helaine
 APPLICANT: Carrer, Helaine
 APPLICANT: Staub, Jeffrey
 APPLICANT: Zoubenko, Oleg V.
 APPLICANT: Allison, Lori A.
 APPLICANT: Kanevski, Ivan
 TITLE OF INVENTION: DNA Constructs and Methods for Stably Transforming Plastids of Multicellular Plants and Expressing Recombinant Proteins Therein
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dann, Dorfman, Herrrell and Skillman
 STREET: 1601 Market Street Suite 720
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103-2307
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/193,853
 FILING DATE: 01-MAY-1990
 CLASSIFICATION:
 PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Janet E.
 REGISTRATION NUMBER: 36,252
 TELEPHONE: (215) 563-4100
 TELEFAX: (215) 563-4044
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 168 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: not relevant
 Molecule type: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-09-193-853-2

Query Match 46.9%; Score 85.4%; DB 2; Length 168;
 Best Local Similarity 98.9%; Pred. No. 3.7e-20;
 Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GCTCCCCGGCTGGTCAATGAGAAATGATAAAGGCTCGGGATGAGCTGGAGGG 66
 Db 1 GCTCCCCGGCTGGTCAATGAGAAATGATAAAGGCTCGGGATGAGCTGGAGGG 60

Qy 67 CAGGGATGGCTATATTCTGGAGGA 93
 Db 61 CAGGGATGGCTATATTCTGGAGGA 87

RESULT 14
 US-09-283-419-3
 Sequence 3, Application US/09283419A
 ; Patent No. 6218145
 ; GENERAL INFORMATION:
 ; APPLICANT: Bogosian, Gregg
 ; ATTORNEY: O'Neil, Julia P.
 ; APPLICANT: Staub, Jeffrey M.
 ; TITLE OF INVENTION: Mitochondrial Promoter Combinations
 ; FILE REFERENCE: MOPV04-
 ; CURRENT APPLICATION NUMBER: US/09/283,419A
 ; CURRENT FILING DATE: 1999-04-01
 ; EARLIER APPLICATION NUMBER: 60/080,432
 ; EARLIER FILING DATE: 1998-04-02
 ; NUMBER OF SEQ ID NOS: 3
 ; SEQ ID NO 3
 ; SOFTWARE: PatentIn Ver. 2.1
 ; LENGTH: 184
 ; TYPE: DNA
 ; ORGANISM: Nicotiana tabacum.
 ; US-09-283-419-3

Query Match 46.9%; Score 85.4%; DB 4; Length 184;
 Best Local Similarity 98.9%; Pred. No. 3.8e-20;
 Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GCTCCCCGGCTGGTCAATGAGAAATGATAAAGGCTCGGGATGAGCTGGAGGG 66
 Db 2 GCTCCCCGGCTGGTCAATGAGAAATGATAAAGGCTCGGGATGAGCTGGAGGG 61

Qy 67 CAGGGATGGCTATATTCTGGAGGA 93
 Db 62 CAGGGATGGCTATATTCTGGAGGA 88

RESULT 15
 US-08-189-256A-24
 Sequence 24, Application US/08189256A
 ; Patent No. 5877402
 ; GENERAL INFORMATION:
 ; APPLICANT: Maliga, Pal
 ; ATTORNEY: Svab, Zora
 ; APPLICANT: Staub, Jeffrey
 ; ATTORNEY: Zoubenko, Oleg V.
 ; ATTORNEY: Allison, Lori A.
 ; ATTORNEY: Carrer, Helaine
 ; TITLE OF INVENTION: DNA Constructs and Methods for Stably Transforming Plastids of Multicellular Plants and Expressing Recombinant Proteins Therein
 ; NUMBER OF SEQUENCES: 47
 ; CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dann, Dorfman, Herrell and Skillman
 STREET: 1601 Market Street Suite 720
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103-2307
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/193,853
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/189,256
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/518,763
 FILING DATE: 01-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Reed, Janet E.
 REGISTRATION NUMBER: 36,252
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 563-4100

```

TITLE OF INVENTION: DNA Constructs and Methods for Stably
TITLE OF INVENTION: Transforming Plastids of Multicellular Plants and
TITLE OF INVENTION: Expressing Recombinant Proteins Therein
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dunn, Dorfman, Herrell and Skillman
STREET: 1601 Market Street Suite 720
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/189,256A
FILING DATE: 31-JAN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/111,398
FILING DATE: 25-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/518,763
FILING DATE: 01-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36,252
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 563-4100
TELEFAX: (215) 563-4044
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 258 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-189-256A-24

```

```

Query Match 46.0% Score 85.4; DB 2; Length 258;
Best Local Similarity 98.9%; Pred. No. 4, 4e-20;
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 7 GCTCCGCCGCGTCGTCAATGAGATGGATAAGGGCTGTGGATTCAGTGAGGGG 66
Db 17 GCTCCGCCGCGTCGTCAATGAGATGGATAAGGGCTGTGGATTCAGTGAGGGG 76
Qy 67 CAGGGATGGATATAATTCTGGAGGA 93
Db 77 CAGGGATGGATATAATTCTGGAGGA 103

```

Search completed: May 25, 2003, 14:27:51
Job time : 62 secs

20	66.8	36.7	1241	9	US-09-987-107-65	
	21	66.8	36.7	1282	9	US-09-987-107-43
	22	66.8	36.7	1285	9	US-09-987-107-45
	23	66.8	36.7	1301	9	US-09-987-107-67
	24	66.8	36.7	4100	9	US-09-813-718-5
	25	66.4	36.5	4682	9	US-09-813-718-3
	26	66.4	36.5	4682	9	US-09-813-718-7
	27	66.4	36.5	5174	9	US-09-813-718-1
	28	65.6	36.0	4877	9	US-09-813-718-11
	29	65.4	35.9	656	9	US-09-976-297-5
	30	65.4	35.9	656	9	US-09-976-297-7
	31	65.4	35.9	4742	9	US-09-813-718-15
	32	65.4	35.9	4811	9	US-09-813-718-13
C	33	64.6	35.5	4974	9	US-09-906-209-17
C	34	64.6	35.5	6611	9	US-09-934-900-26
	35	62	34.1	5770	10	US-09-838-718A-7
	36	62	34.1	5870	10	US-09-838-718A-8
	37	62	34.1	5906	10	US-09-838-718A-6
	38	60.4	33.2	1327	9	US-09-940-235-6
	39	60.4	33.2	1377	9	US-09-940-235-5
	40	60.4	33.2	1661	9	US-09-940-235-10
	41	60.4	33.2	1782	9	US-09-940-235-11
	42	60.4	33.2	2096	9	US-09-940-235-12
	43	59.6	32.7	5230	10	US-09-838-718A-5
	44	59.4	32.6	6714	9	US-09-981-002-5
	45	58.4	32.1	682	9	US-10-212-357-1

RESULT 2

US-10-109-812-1
 ; Sequence 1, Application US/10109812
 ; Publication No. US20030088081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Maligna, Pal

Qy 1 GAGCTCGCTCCCGCCCGTCAATGAGAATGGATAAGAGGCTCGTGGGATTGACCTG 60
 Db 1 GAGCTCGCTCCCGCCCGTCAATGAGAATGGATAAGAGGCTCGTGGGATTGACCTG 60

RESULT 4

US-09-940-925A-163
 ; Sequence 163, Application US/0940925A
 ; Publication No. US20030054338A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BROW, MARY ANN D.
 ; LIAMICHEN, VICTOR I.
 ; OLIVE, DAVID M.

Qy 61 AGGGGGCAGGATGGCTATATTCTGGGAG 90
 Db 61 AGGGGGCAGGATGGCTATATTCTGGGAG 90

RESULT 4

US-09-940-925A-163
 ; Sequence 163, Application US/0940925A
 ; Publication No. US20030054338A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MEDDLEN, CARROLL
 ; LIAMICHEN, VICTOR I.
 ; OLIVE, DAVID M.

Qy 165 TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF PATHOGENS
 ; NUMBER OF SEQUENCES: 165
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MEDDLEN, CARROLL
 ; STREET: 220 MONTGOMERY STREET, SUITE 2200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 94104

Qy 165 COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; CURRENT APPLICATION NUMBER: #1.0, Version #1.30
 ; SOFTWARE: Patientin Release #1.0, Version #1.30
 ; CURRENT APPLICATION NUMBER: US/09/940, 925A
 ; APPLICATION NUMBER: US-09/940, 925A
 ; FILING DATE: 10-Jun-2002
 ; CLASSIFICATION: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CARROLL, PETER G.
 ; REGISTRATION NUMBER: 32,637
 ; REFERENCE/DOCKET NUMBER: FORS-017556
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705 8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 163:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 144 base pairs
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 163:
 ; US-09-940-925A-163

Query Match 49.5%; Score 90; DB 9; Length 130;
 Best Local Similarity 100.0%; Prid. No. 2.1e-20;
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGCTCGCTCCCGCCCGTCAATGAGAATGGATAAGAGGCTCGTGGGATTGACCTG 60
 Db 1 GAGCTCGCTCCCGCCCGTCAATGAGAATGGATAAGAGGCTCGTGGGATTGACCTG 60

Qy 61 AGGGGGCAGGATGGCTATATTCTGGGAG 90
 Db 61 AGGGGGCAGGATGGCTATATTCTGGGAG 90

RESULT 3

US-10-109-812-4
 ; Sequence 4, Application US/10109812
 ; Publication No. US20030088081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Maligna, Pal
 ; APPLICANT: Cornelle, Sylvie

Qy 163 TITLE OF INVENTION: High Level Expression of Immunogenic Proteins in the Plastics of Higher Plants
 ; TITLE OF INVENTION: Plastics of Higher Plants
 ; FILE REFERENCE: Rutgers-00-0038 CIP
 ; CURRENT APPLICATION NUMBER: US/10/109, 812
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: PCT/US00/25930
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: 60/211, 139
 ; PRIOR FILING DATE: 2000-06-13
 ; PRIOR APPLICATION NUMBER: 60/155, 007
 ; PRIOR FILING DATE: 1999-09-21
 ; NUMBER OF SEQ ID NOS: 57
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 4
 ; LENGTH: 131
 ; TYPE: DNA
 ; ORGANISM: tobacco

Qy 90 AGGGAGCCACCAAGGCTTCCCACTAGAAATATTGTTAACCTAAAGGAGAT 148
 Db 37 AGGGAGCCACCAAGGCTTCCCACTAGAAATATTGTTAACCTAAAGGAGAT 96

Qy 149 ACATAGGCAAGCATGACTGGGACAG 176
 Db 97 ACATAGGCAAGCATGACTGGGACAG 124

RESULT 5

US-09-843-324A-1
 ; Sequence 1, Application US/09843324A
 ; Patient No. US2002004293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Staub, Jeffrey
 ; APPLICANT: Ye, Guangning

Qy 131 Query Match 49.5%; Score 90; DB 9; Length 131;
 Best Local Similarity 100.0%; Prid. No. 2.1e-20; Mismatches 0; Indels 0; Gaps 0;

Qy 90 Query Match 49.5%; Score 90; DB 9; Length 131;
 Best Local Similarity 99.9%; Prid. No. 2.1e-20; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
US-09-843-324A-2
Sequence 2, Application US/09843324A
Patent No. US2002004934A1
GENERAL INFORMATION:
APPLICANT: Staub, Jeffrey
APPLICANT: Ye, Guangning
APPLICANT: Broyles, Debra
TITLE OF INVENTION: Method for the transformation of plant cell plantan

RESULT 7
US-10-109-812-41
; Sequence 41, Application US/10109812
; Publication No. US20030088081A1
; GENERAL INFORMATION:
; APPLICANT: Malica Pa

RESULT 9
 US-09-897-776A-17
 ; Sequence 17, Application US/09897776A
 ; PUBLICATION NO. US2003-0092001A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Feng
 ; APPLICANT: Schnable, Patrick S.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING MULTIPLE
 ; TITLE OF INVENTION: START CODONS AND HISTIDINE TAGS
 ; FILE REFERENCE: 08411-027001
 ; CURRENT APPLICATION NUMBER: US/09/897,776A
 ; CURRENT FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: US 09/7732, 9,90
 ; PRIOR FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: US 60/169,725
 ; PRIOR FILING DATE: 1999-12-08
 ; NUMBER OF SEQ ID NOS: 37
 ; SEQ ID NO: 17
 ; LENGTH: 97
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetically generated oligonucleotide
 ; US-09-897-776A-17

Query Match Score 68.8; DB 9; Length 97;
 Best Local Similarity 97.2%; Pred. No. 2, 2e-13;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 89 AGGGAGACACACAGGGTTCCCACTAGAAATAATTGGTTAACCTTAAGGAGAT 148
 Db 17 AGGGAGACACACAGGGTTCCCACTAGAAATAATTGGTTAACCTTAAGGAGAT 76

Qy 149 ACATATGGCAAG 160
 Db 77 ACATATGGCATG 88

RESULT 10
 US-09-813-718-9
 ; Sequence 9, Application US/09813718
 ; PUBLICATION NO. US2002-0182666A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schimmei, Paul
 ; APPLICANT: Wakasugi, Keisuke
 ; TITLE OF INVENTION: Human Aminocetyl-tRNA Synthetase Polypeptides Useful For
 ; THE REGULATION OF ANGIogenesis
 ; FILE REFERENCE: 00-221
 ; CURRENT APPLICATION NUMBER: US/09/813,718
 ; CURRENT FILING DATE: 2001-03-21
 ; NUMBER OF SEQ ID NOS: 58
 ; SEQ ID NO: 9
 ; LENGTH: 5018
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (34-58)..(4879)
 ; OTHER INFORMATION: Description of Artificial Sequence: human
 ; OTHER INFORMATION: full-length trpRS in PET20B

Qy 89 AGGGAGACACACAGGGTTCCCACTAGAAATAATTGGTTAACCTTAAGGAGAT 148
 ; US-09-813-718-9

Query Match Score 67.6; DB 9; Length 5018;
 Best Local Similarity 94.6%; Pred. No. 3, 3e-12;
 Matches 70; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 89 AGGGAGACACACAGGGTTCCCACTAGAAATAATTGGTTAACCTTAAGGAGAT 148
 ; US-09-813-718-9

RESULT 11
 US-09-987-107-51
 ; Sequence 51, Application US/09987107
 ; PATENT NO. US2002-0156007A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRAVERSSEN, Jonas
 ; APPLICANT: MOESTRUP, Soren
 ; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
 ; FILE REFERENCE: GRAVERSSEN
 ; CURRENT APPLICATION NUMBER: US/09/987,107
 ; CURRENT FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/264,022
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: DK PA2001 00057
 ; PRIOR FILING DATE: 2001-01-15
 ; PRIOR APPLICATION NUMBER: DK PA2000 01682
 ; PRIOR FILING DATE: 2000-11-10
 ; NUMBER OF SEQ ID NOS: 91
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO: 51
 ; LENGTH: 1057
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: PT7 H6 Fx Cys-Apo A1 Plasmid
 ; NAME/KEY: CDS
 ; LOCATION: (100)..(882)
 ; OTHER INFORMATION:
 ; US-09-987-107-51

Query Match Score 36.7%; DB 9; Length 1057;
 Best Local Similarity 97.1%; Pred. No. 3e-12;
 Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 89 AGGGAGACACACAGGGTTCCCACTAGAAATAATTGGTTAACCTTAAGGAGAT 148
 Db 36 AGGGAGACACACAGGGTTCCCACTAGAAATAATTGGTTAACCTTAAGGAGAT 95

RESULT 12
 US-09-987-107-49
 ; Sequence 49, Application US/09987107
 ; PATENT NO. US2002-0156007A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GRAVERSSEN
 ; APPLICANT: MOESTRUP, Soren
 ; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
 ; FILE REFERENCE: GRAVERSSEN
 ; CURRENT APPLICATION NUMBER: US/09/987,107
 ; CURRENT FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/264,022
 ; PRIOR FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: DK PA2001 00057
 ; PRIOR FILING DATE: 2001-01-15
 ; PRIOR APPLICATION NUMBER: DK PA2000 01682
 ; NUMBER OF SEQ ID NOS: 91
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO: 49
 ; LENGTH: 1086
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence

FEATURE:
; OTHER INFORMATION: PT7H6 Trip-A-Apo A-1-del 43 - AmpR plasmid
; NAME/KEY: CDS
; LOCATION: (100)..(918)
; OTHER INFORMATION: US-09-987-107-49

Query Match 36.7%; Score 66.8; DB 9; Length 1088;
Best Local Similarity 97.1%; Pred. No. 3.1e-12; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
SEQ ID NO: 53

RESULT 13
US-09-987-107-47
; Sequence 47, Application US/09987107
; Patent No. US20020156007A1

GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin version 3.1
; SEQ ID NO: 47
; LENGTH: 1217

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: PT7 H6 Trip-A-Apo A-1 - AmpR plasmid
; NAME/KEY: CDS
; LOCATION: (100)..(1047)
; OTHER INFORMATION: US-09-987-107-47

Query Match 36.7%; Score 66.8; DB 9; Length 1217;
Best Local Similarity 97.1%; Pred. No. 3.2e-12; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
SEQ ID NO: 55

RESULT 14
US-09-987-107-53
; Sequence 53, Application US/09987107
; Patent No. US20020156007A1

GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin version 3.1

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: PT7H6 Trip-A-Fn-Apo A1 - AmpR plasmid
; NAME/KEY: CDS
; LOCATION: (100)..(1068)
; OTHER INFORMATION: US-09-987-107-53

Query Match 36.7%; Score 66.8; DB 9; Length 1235;
Best Local Similarity 97.1%; Pred. No. 3.2e-12; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
SEQ ID NO: 57

RESULT 15
US-09-987-107-55
; Sequence 55, Application US/09987107
; Patent No. US20020156007A1

GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin version 3.1

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: PT7H6 Trip-A-Fn-Apo A1 - AmpR plasmid
; NAME/KEY: CDS
; LOCATION: (100)..(1068)
; OTHER INFORMATION: US-09-987-107-55

Query Match 36.7%; Score 66.8; DB 9; Length 1235;
Best Local Similarity 97.1%; Pred. No. 3.2e-12; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
SEQ ID NO: 59

RESULT 16
US-09-987-107-57
; Sequence 57, Application US/09987107
; Patent No. US20020156007A1

GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin version 3.1

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: PT7H6 Trip-A-Fn-Apo A1 - AmpR plasmid
; NAME/KEY: CDS
; LOCATION: (100)..(1068)
; OTHER INFORMATION: US-09-987-107-57

Query Match 36.7%; Score 66.8; DB 9; Length 1235;
Best Local Similarity 97.1%; Pred. No. 3.2e-12; Indels 0; Gaps 0;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
SEQ ID NO: 61

Tue May 27 10:47:40 2003

us-09-762-105a-14.rnpb

Page 6

Db 96 ACATATGGAA 105

Search completed: May 25, 2003, 14:29:51
Job time : 107 secs

GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 13:23:39 ; Search time 1416 Seconds

(without alignments)
2081.623 Million cell updates/sec

Title: US-09-762-105A-14

Perfect score: 182

Sequence: 1 gagctcgcccccggcgtc.....tgacttgtggacaggcttagc 182

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: em_estba:*

2: em_estbh:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estp1:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_eston:*

17: em_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_p1n:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

RESULT 1

BH58942

LOCUS

BOHLZ7TR

Brassica oleracea genomic clone BOHLZ72, DNA sequence.

DEFINITION

BH58942

ACCESSION

BH58942.1

VERSION

1

KEYWORDS

GSS.

ORGANISM

Brassica oleracea.

REFERENCE

1 (bases 1 to 787)

AUTHORS

Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.

TITLE

Whole genome shotgun sequencing of Brassica oleracea

JOURNAL

unpublished (2001)

COMMENT

Other.GSS: BOHLZ7TR

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtoinet@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1

Source

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

1

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: "TF
 Class: sheared ends.
 Location/Qualifiers
 1. "organism="Brassica oleracea"
 /strain="TO1000D3"
 /db_xref="taxon:3712"
 /clone_id="BOGI"
 /note="Vector: pHO51; Site_1: BstX1; 2-3 kb sheared
 genomic DNA inserted into pHO51 using BstX1 linkers"
 BASE COUNT 233 a 142 c 191 g 272 t
 ORIGIN
 Query Match 39.7%; Score 72.2; DB 17; Length 838;
 Best Local Similarity 80.0%; Pred. No. 2.9e-12;
 Matches 80; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 ORGANISM
 Qy 4 ctcctcccccgcgcgtcgatcaatggatggataaggctcggtggatgtggatgtgg 63
 Db 720 cttgcctccctcgatgtggatgtggataaggctcggtggatgtggatgtgg 79
 Qy 64 GGGCAGGGATGGCTATATTCTGGGAGGGAGC 96
 Db 780 GGGTAGGGTAGCTATATTCTGGGAGGCAAC 812
 RESULT 5
 BH668002 BH668002 770 bp DNA linear GSS 19-FEB-2002
 LOCUS BOMN95TF BO_2_3_KB Brassica oleracea genomic clone BOMN95, DNA
 DEFINITION sequence.
 ACCESSION BH668002.1 GI:18727985
 VERSION 1
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM
 Qy 1. 836
 Db 664 CTTGCGTCCCTCGTGTGATGTAAGATGATAAGCTCGTGGATGTGGAGG 723
 Qy 4 CTCGGTCCCGCCCGTCAATGAGCTCGTGGATGTGGAGCTCGTGGAGG 63
 Db 664 CTTGCGTCCCTCGTGTGATGTAAGATGATAAGCTCGTGGATGTGGAGG 723
 Qy 64 GGGCAGGGATGGCTATATTCTGGGAGGGAGCAGCGTTCACCTAGAAATAATT 123
 Db 724 GGTAAGGGTAGCTATATTCTGGGAGGCACTGGGAGCTCGTGGATGTGGAGG 783
 Qy 124 TTGTGTTAACTTTAAGAGGATATA 150
 Db 784 AAGTATGACTTGGAAAATGAGAAATT 810
 RESULT 7
 BH653765 BH653765 841 bp DNA linear GSS 19-FEB-2002
 LOCUS BOMN13TF BO_2_3_KB Brassica oleracea genomic clone BOMN13, DNA
 DEFINITION sequence.
 ACCESSION BH653765

FEATURES source
 1. "organism="Brassica oleracea"
 /strain="TO1000D3"
 /db_xref="taxon:3712"
 /clone_id="BO_2_3_KB"
 /note="Vector: pHO51; Site_1: BstX1; 2-3 kb sheared
 genomic DNA inserted into pHO51 using BstX1 linkers"
 BASE COUNT 204 a 162 c 246 g 158 t
 ORIGIN
 Query Match 39.5%; Score 71.8; DB 17; Length 770;
 Best Local Similarity 86.8%; Pred. No. 3.8e-12;

VERSION BH653765.1 GI:18711917
 SOURCE GSS. *Brassica oleracea*.
 ORGANISM *Brassica oleracea*; *Viridiplantae*; *Streptophyta*; *Tracheophyta*; *Eukarya*; *Unikonta*; *Amoebozoa*; *Sporomophyta*; *Magnoliophyta*; *Eudicots*; *Rosidae*; *eurosid II*; *Brassicaceae*; *Brassica*.
 REFERENCE 1 (bases 1 to 841)
 AUTHORS Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
 TITLE whole genome shotgun sequencing of *Brassica oleracea*
 JOURNAL unpublished (2001)
 COMMENT Contact: Chris Town
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 /db_xref="TIGR09712"
 Class: sheared ends.
 FEATURES SOURCE 1. 841
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /clone="BOMN13"
 /note="Vector: PHOS1; Site_1: BstX1; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstX1 linkers"
 BASE COUNT 236 a 146 c 194 g 265 t
 ORIGIN Query Match 39.5%; Score 71.8; DB 17; Length 841;
 Best Local Similarity 68.0%; Pred. No. 3.9e-12;
 Matches 100; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
 QY 4 CTCGAGGGATGGCTATATTCTGGAGGGACCAACAGCTTCCCACTAGAAATAATT 123
 DB 675 CTCGAGGGATGGCTATATTCTGGAGGGACCAACAGCTTCCCACTAGAAATAATT 123
 QY 64 GGGCAGGGATGGCTATATTCTGGAGGGACCAACAGCTTCCCACTAGAAATAATT 123
 DB 735 GGGTAGGGATGGCTATATTCTGGAGGGACCAACAGCTTCCCACTAGAAATAATT 123
 QY 124 TGTGTTAACTTAAAGAGGAGATATAC 150
 DB 795 AAGTATGACTTGGAAATGAGAACATC 821

RESULT 9 BH474481
 LOCUS BH474481
 DEFINITION BOHY09TR BOHH *Brassica oleracea* genomic clone BOHY09, DNA sequence.
 ACCESSION BH474481
 VERSION BH474481.1 GI:17682592
 KEYWORDS GSS.
 SOURCE *Brassica oleracea*.
 ORGANISM *Brassica oleracea*; *Viridiplantae*; *Streptophyta*; *Tracheophyta*; *Eukarya*; *Unikonta*; *Amoebozoa*; *Sporomophyta*; *Magnoliophyta*; *Eudicots*; *Rosidae*; *eurosid II*; *Brassicaceae*; *Brassica*.
 REFERENCE 1 (bases 1 to 805)
 AUTHORS Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of *Brassica oleracea*
 JOURNAL Unpublished (2001)
 COMMENT Contact: Chris Town
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-3523
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 /db_xref="TIGR09712"
 Class: sheared ends.
 FEATURES SOURCE 1. 805
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /db_xref="TIGR09712"
 /clone="BOHH"
 /note="Vector: PHOS1; Site_1: BstX1; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstX1 linkers"
 BASE COUNT 221 a 139 c 182 g 263 t
 ORIGIN Query Match 39.0%; Score 71; DB 17; Length 805;
 Best Local Similarity 84.2%; Pred. No. 7e-12;
 Matches 80; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 4 CTCGCTCCCGCGCTGCGTCAATGAGATGGATAACAGGCGTGGATGACGTGG 63

Db	697	CTTGGTCCCCCGCTGTGATCGAATAAGAATGGATAAGGGCTGGGATGTGACTGTGACGG 756	AUTHORS	Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterback,T.,
Qy	64	GGCAGGGATGGCTATATTCCTGGAGGAGACCA 98	TITLE	Feldblyum,T., Liang,F., Creasy,T., and Fraser,C.M.
Db	757	GGCTAGGGCTAGCTATATTCCTGGAGGCAACTCA 791	JOURNAL	Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
COMMENT		unpublished (2000)	COMMENT	Xiaoying Lin
RESULT 10			COMMENT	The Institute for Genomic Research
BH721450			COMMENT	9712 Medical Center Dr., Rockville, MD 20850, USA
LOCUS	BH721450	847 bp DNA linear GSS 20-FEB-2002	FEATURES	Class: shotgun
DEFINITION	BOMFR11TR BO_2_3_KB	Brassica oleracea genomic clone BOMFR11, DNA sequence.	source	Location/Qualifiers
REFERENCE	BH721450	1 (bases 1 to 847)		1..169
AUTHORS	Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.			/organism="Arabidopsis thaliana"
TITLE	Whole genome shotgun sequencing of Brassica Oleracea			/strain="Landsberg erecta"
JOURNAL	Unpublished (2001)			/db_xref="3702"
COMMENT	Other_GSS: BOMFR11TR			/clone="LERG142"
SOURCE	Brassica oleracea.			/clone_lib="LERG"
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.			/note="Organ: Leaf; Vector: PUC19JK; Total genomic DNA was sheared to 0.4-0.7 kbp before ligation."
FEATURES	Brassica oleracea.			
source				
BASE COUNT				46 a 49 c 27 g 47 t
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 70.8%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 0; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 38.9%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 39.0%; DB 17; Length 847;
source				Best Local Similarity 84.2%; Pred. No. 7.1e-12;
FEATURES				Mismatches 0; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches 12; Indels 0; Gaps 0;
source				
BASE COUNT				
ORIGIN				
Query Match				Query Match
Best Local Similarity				Best Local Similarity
Matches				Score 46.1%; DB 17; Length 169;
source				Best Local Similarity 86.7%; Pred. No. 5e-12;
FEATURES				Mismatches

Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: *TF*

Class: sheared ends.

FEATURES

source

1. 305
 /organism="Brassica oleracea"
 /strain="T01000D3"
 /db_xref="taxon:312"
 /clone="BOGN48"
 /clone.lib="BO_2.3_KB"
 /note="Vector: phoS1; Site_1: BstXI; 2-3 kb sheared
 genomic DNA inserted into phoS1 using BstXI linkers"
 88 a 85 c 57 g 75 t

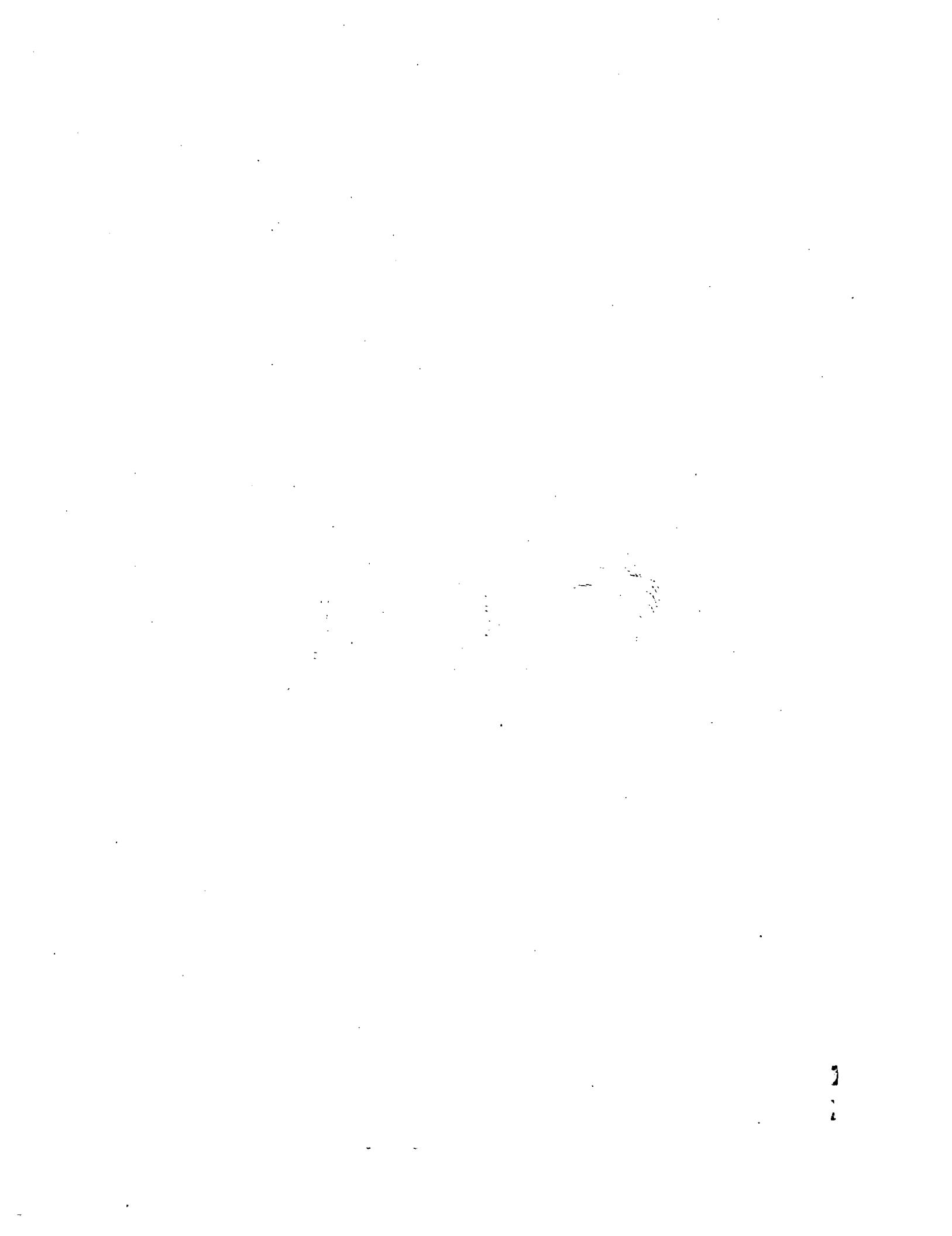
BASE COUNT

ORIGIN

	Query Match	Score	DB	Length
Qy	CTCGCTCCCCGGCGGTGTCATTGAGATGGATAAGGGCTCTGGATGACGTGAGG	38 9%	17	305
Db	CTTCTCCCTGCCTGATGATGATAAGGGCTGGATGACGTGAGG	86 7%	12	6-12;
Qy	GGCAGGGATGGCTATATTCTGGAGGA	4	11	93
Db	GGGTAGGGTAGCTATTCCTGGAGCA	42	11	13

Best Local Similarity
 Matches 78; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Score completed: May 25, 2003, 14:26:40
 Job time : 1420 secs



GenCore version 5.1.5
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 12:31:19 ; Search time 210 Seconds

(without alignments)
1951.732 Million cell updates/sec

Title: US-09-762-105A-14

Perfect score: 182
Sequence: 1 gagctcgctcccgccgtc.....tgactgtggacaggcttagc 182

Scoring table: IDENTITY_NUC

GapOp 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Total number of hits satisfying chosen parameters: 4370478

RESULT 1
ID AAZ61373 standard; DNA; 182 BP.

XX AC AAZ61373;

XX DT 19-JUN-2000 (first entry)

XX DE Nucleotide sequence of chimeric promoter PcrnLrt7g10+DB/EC.

XX KW Chimeric promoter; plasmid rRNA operon omega-type promoter; haemoglobin; Prrn promoter; protein expression; vaccine; enzyme; T7 phage gene 10; ss.

XX OS Synthetic.

XX FH Key

FT promoter 7..89

FT /tag-^a

FT misc_signal 14..144

FT /tag-^b

FT /note- "Shine-Dalgarno sequence"

Nucleotide sequence	10	93.6	51.4	227	21	AAZ61360
DNA fragment design	11	93.6	51.4	227	21	AAZ61362
Nucleotide sequence	12	93.6	51.4	1049	22	AAF81268
Nucleotide sequence	13	93.6	51.4	1985	21	AAZ61382
Nucleotide sequence	14	92.2	50.7	153	21	AAZ61369
Nucleotide sequence	15	92.2	50.7	185	21	AAZ61372
Nucleotide sequence	16	92.2	50.7	195	21	AAZ61368
Plasmid rRNA operon	17	90	49.5	130	22	AAFB1256
Plasmid rRNA operon	18	90	49.5	131	22	AAF81159
Nucleotide sequence	19	90	49.5	154	21	AAZ61364
Nucleotide sequence	20	90	49.5	159	21	AAZ61366
Nucleotide sequence	21	90	49.5	195	21	AAZ61365
Nucleotide sequence	22	90	49.5	195	21	AAZ61367
Nucleotide sequence	23	90	49.5	196	21	AAZ61363
Nucleotide sequence	24	90	49.5	1953	21	AAZ61381
PrnLact1 Spns Pro	25	89.6	49.2	171	16	AAQ94853
Nucleotide sequence	26	89	48.9	183	21	AAZ61371
Nucleotide sequence	27	89	48.9	201	21	AAZ61370
Regulatory region	28	86.4	47.5	129	20	AAZ21430
5', and 3', regulatory	29	86	47.3	140	20	AAZ21424
Regulatory region	30	86	47.3	164	20	AAZ21431
Plasmid 16S rRNA	31	85.8	47.1	127	22	AAFB81276
DNA of upstream re	32	85.8	47.1	201	18	AAAT66301
Chloroplast gene t	33	85.8	47.1	6477	22	AAF16797
5', and 3', regulato	34	85.4	46.9	161	20	AAZ21423
Prn promoter sequ	35	85.4	46.9	165	20	AAZ21409
Prn/rbcL/Rubisco	36	85.4	46.9	168	20	AAZ21407
Tobacco Plastid RR	37	85.4	46.9	184	22	AAFB8143
Regulatory region	38	85.4	46.9	258	20	AAZ21429
Plastid targeting	39	85.4	46.9	1134	18	AAAT85195
Selectable marker	40	85.4	46.9	163	20	AAZ21415
Nucleotide sequence	41	85.4	46.9	1183	21	AAZ61376
Regulatory region	42	85.4	46.9	1208	20	AAZ21433
Regulatory region	43	85.4	46.9	1416	18	AAZ21432
Plastid targeting	44	85.4	46.9	1417	18	AAAT85196
targetting region	45	85.4	46.9	2962	20	AAZ21408

ALIGNMENTS

Nucleotide sequence	10	93.6	51.4	227	21	AAZ61360
DNA fragment design	11	93.6	51.4	227	21	AAZ61362
Nucleotide sequence	12	93.6	51.4	1049	22	AAF81268
Nucleotide sequence	13	93.6	51.4	1985	21	AAZ61382
Nucleotide sequence	14	92.2	50.7	153	21	AAZ61369
Nucleotide sequence	15	92.2	50.7	185	21	AAZ61372
Nucleotide sequence	16	92.2	50.7	195	21	AAZ61368
Plasmid rRNA operon	17	90	49.5	130	22	AAFB1256
Plasmid rRNA operon	18	90	49.5	131	22	AAF81159
Nucleotide sequence	19	90	49.5	154	21	AAZ61364
Nucleotide sequence	20	90	49.5	159	21	AAZ61366
Nucleotide sequence	21	90	49.5	195	21	AAZ61365
Nucleotide sequence	22	90	49.5	195	21	AAZ61367
Nucleotide sequence	23	90	49.5	196	21	AAZ61363
Nucleotide sequence	24	90	49.5	1953	21	AAZ61381
PrnLact1 Spns Pro	25	89.6	49.2	171	16	AAQ94853
Nucleotide sequence	26	89	48.9	183	21	AAZ61371
Nucleotide sequence	27	89	48.9	201	21	AAZ61370
Regulatory region	28	86.4	47.5	129	20	AAZ21430
5', and 3', regulato	29	86	47.3	140	20	AAZ21424
Regulatory region	30	86	47.3	164	20	AAZ21431
Plasmid 16S rRNA	31	85.8	47.1	127	22	AAFB81276
DNA of upstream re	32	85.8	47.1	201	18	AAAT66301
Chloroplast gene t	33	85.8	47.1	6477	22	AAF16797
5', and 3', regulato	34	85.4	46.9	161	20	AAZ21423
Prn promoter sequ	35	85.4	46.9	165	20	AAZ21409
Prn/rbcL/Rubisco	36	85.4	46.9	168	20	AAZ21407
Tobacco Plastid RR	37	85.4	46.9	184	22	AAFB8143
Regulatory region	38	85.4	46.9	258	20	AAZ21429
Plastid targeting	39	85.4	46.9	1134	18	AAAT85195
Selectable marker	40	85.4	46.9	163	20	AAZ21415
Nucleotide sequence	41	85.4	46.9	1183	21	AAZ61376
Regulatory region	42	85.4	46.9	1208	20	AAZ21433
Regulatory region	43	85.4	46.9	1416	18	AAZ21432
Plastid targeting	44	85.4	46.9	1417	18	AAAT85196
targetting region	45	85.4	46.9	2962	20	AAZ21408

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description	
1	182	100.0	182	21	AAZ61373	Nucleotide sequence	
2	182	100.0	1961	21	AAZ61384	Nucleotide sequence	
C	3	182	100.0	5263	21	AAZ61386	Nucleotide sequence
4	164.4	90.3	182	21	AAZ61374	Nucleotide sequence	
5	159.4	87.6	161	21	AAZ61375	Nucleotide sequence	
6	111.2	61.1	168	21	AAZ61376	Prn/G10 fusion n	
7	111.2	61.1	168	22	AAF5352	Nucleotide sequence	
8	111.2	61.1	168	22	AAF7902	Prn/G10L promoter	
9	93.6	51.4	191	21	AAZ61361	Nucleotide sequence	

Key	Location/Qualifiers
FFI	7..176
FTT	/*tag= a
FTT	/*note= "T7 phage gene 10 downstream box"
FTT	218..1001
FTT	/*tag= "b"
FTT	/*note= "addA sequence"
FTT	1035..1755
FTT	/*tag= "c"
FTT	/*note= "green fluorescent protein region"
FTT	1766..1954
FTT	/*tag= "d"
FTT	/*note= "pbA region"
XX	
PN	W020007431-A1.
XX	
XX	17-FEB-2000.
PD	
XX	03-AUG-1999;
XX	99WO-US17806.
PR	03-AUG-1998;
PR	98US-0095161.
PR	03-AUG-1998;
PR	98US-0095167.
PR	15-DEC-1998;
PR	98US-0112257.
PR	29-APR-1999;
PR	99US-0131611.
PR	11-JUN-1999;
PR	99US-0138766.
XX	
(RUTTF) UNIV RUTGERS STATE NEW JERSEY .	
XX	
XX	Maliga P, Kuroda H, Khan MS;
PI	
XX	
DR	2000-20525/18.
PT	
PT	New recombinant DNA constructs, for expressing high level heterologous protein in plastids of higher plants, includ-
PT	leader sequence and a downstream box element -
PT	
XX	
PS	Disclosure: Fig 32: 164pp; English.
XX	
CC	The present sequence represents a DNA construct of the in-
CC	CC heterologous proteins in the plastids of higher plants. The
CC	CC constructs comprise a 5' regulatory region which includes
CC	CC a leader sequence and a downstream box element or
CC	CC element, a leader sequence and a downstream box element or
CC	CC a coding region of the heterologous protein. The chime
CC	CC region enhances translational efficiency of an mRNA molecule
CC	CC the DNA construct. The DNA constructs are used for produc-
CC	CC tion of monocot and dicot plants having high levels of heterologous
CC	CC expression. The DNA constructs can be used to drive expression of prote-
CC	CC in agronomic, industrial or pharmaceutical importance, includ-
CC	CC ing vaccines, healthcare products like human haemoglobin,
CC	CC household enzymes. Plants which can be transformed with the
CC	CC of the invention include maize, millet, sorghum, sugar can-
CC	CC wheat, barley, oat, rye or turf grass.
XX	
SQ	Sequence 1961 BP: 552 A; 430 C; 496 G; 493 T; 0 other;
Query	Query Match 100.0%; Score 182; DB 21; Length 182; Conservative 0; Mismatches 0; Indels 0
Db	1 GAGCTGCTCCCCCGCGCTTCAATGAGAATGGATAGGCTCTGGGG
Db	1 GAGCTGCTCCCCGGCTTCAATGAGAATGGATAGGCTCTGGGG
Qy	61 AGGGGGCAGGGATGGCTATATTCTGGGGAGACCCACACGGTTTCCCA
Db	61 AGGGGGCAGGGATGGCTATTTCTGGGGAGACCCACACGGTTTCCCA
Qy	.121 ATTTGGTTAATCTTAAGAAGGAGATATACTATGGCAGCATACAGTGG
Db	.121 ATTTGGTTAATCTTAAGAAGGAGATATACTATGGCAGCATACGTGG

FT misc_signal 114..116 /*tag-^b *note- "Shine-Dalgarno sequence"

FT XX WO200007431-A1.

PN XX

PD 17-FEB-2000.

XX

PF 03-AUG-1999; 99WO-US17806.

XX

PR 03-AUG-1998; 98US-0095163.

PR 03-AUG-1998; 98US-0095167.

PR 15-DEC-1998; 98US-0112257.

PR 29-APR-1999; 99US-0131611.

PR 11-JUN-1999; 99US-0138764.

XX

PA (RUTF) UNIV RUTGERS STATE NEW JERSEY.

XX

PI Maliga P, Kuroda H, Khan MS;

XX

DR WPI; 2000-205525/18.

XX

PT New recombinant DNA constructs, for expressing high levels of heterologous protein in plastids of higher plants, includes promoter, a leader sequence and a downstream box element

PT

XX

PS Claim 3; Fig 3A; 164pp; English.

XX

The present sequence represents a chimeric plastid rRNA operon CC omega-type (Prrn) promoter with atpB translation control sequences and CC a wild type downstream box. The chimeric promoter is used, as a 5', CC regulatory sequence, to produce recombinant DNA constructs for expressing heterologous proteins in the plastids of higher plants. The DNA CC constructs comprise a 5' regulatory region which includes a promoter CC element, a leader sequence and a downstream box element operably linked CC to a coding region of the heterologous protein. The chimeric regulatory CC region enhances translational efficiency of an mRNA molecule encoded by CC the DNA construct. The DNA constructs are used for producing transformed CC monocot and dicot plants having high levels of heterologous protein CC expression. They can be used to drive expression of proteins with CC agronomic, industrial or pharmaceutical importance, including production CC of vaccines, healthcare products like human haemoglobin, industrial or CC household enzymes. Plants which can be transformed with the constructs CC of the invention include maize, millet, sorghum, sugar cane, rice, CC wheat, barley, oat, rye or turf grass.

XX

SQ Sequence 227 BP; 58 A; 42 C; 55 G; 72 T; 0 other;

Query Match 51.4%; Score 93.6; DB 21; Length 227;

Best Local Similarity 75.0%; Pred. No. 3e-21;

Matches 117; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 1 GAGGTGCTCCGGCGCTGGTCAATGAGAATGATGGAGCTGGGATTGACGTG 60

Db 1 GAGGTGCTCCGGCGCTGGTCAATGAGAATGATGGAGCTGGGATTGACGTG 60

Qy 61 AGGGGCCAGGGATGGCTATTTCTGGGAGGGACCAACGGTTCCCACTAGAAATA 120

Db 61 AGGGGCCAGGGATGGCTATTTCTGGGAGAACTAACCGATGTCAGGGACATT 120

Qy 121 ATTTGTTAACTTAACGGGATATACTATGG 156

Db 121 TATTTAACATTGATAATTGCAAAACATTTC 156

RESULT 11

ID AA211362 standard; DNA; 227 BP.

XX

AC AA211362;

XX

DT 19-JUN-2000 (first entry)

DE Nucleotide sequence of chimeric promoter PrrnLatpB+DBm.

XX Chimeric promoter; Plastid rRNA operon omega-type promoter; haemoglobin;

KW Prrn promoter; atpB; protein expression; vaccine; enzyme; ss.

XX Synthetic.

OS

FH Key promoter

FT Location/Qualifiers 7..89

FT /*tag-^a *note- "Prrn Plastid promoter"

FT misc_signal 114..116

FT /*tag-^b *note- "Shine-Dalgarno sequence"

XX WO200007431-A1.

XX

PD 17-FEB-2000.

XX

PF 03-AUG-1999;

XX

PR 03-AUG-1998; 98US-0095163.

PR 03-AUG-1998; 98US-0095167.

PR 15-DEC-1998; 98US-0112257.

PR 29-APR-1999; 99US-0131611.

PR 11-JUN-1999; 99US-0138764.

XX

PA (RUTF) UNIV RUTGERS STATE NEW JERSEY.

XX

PI Maliga P, Kuroda H, Khan MS;

XX

DR WPI; 2000-205525/18.

XX

PT New recombinant DNA constructs, for expressing high levels of heterologous protein in plastids of higher plants, includes promoter, a leader sequence and a downstream box element

PT

XX

PS Claim 3; Fig 3A; 164pp; English.

XX

The present sequence represents a chimeric plastid rRNA operon CC omega-type (Prrn) promoter with atpB translation control sequences and CC a mutated downstream box. The chimeric promoter is used, as a 5', CC regulatory sequence, to produce recombinant DNA constructs for expressing heterologous proteins in the plastids of higher plants. The DNA CC constructs comprise a 5' regulatory region which includes a promoter CC element, a leader sequence and a downstream box element operably linked CC to a coding region of the heterologous protein. The chimeric regulatory CC region enhances translational efficiency of an mRNA molecule encoded by CC the DNA construct. The DNA constructs are used for producing transformed CC monocot and dicot plants having high levels of heterologous protein CC expression. They can be used to drive expression of proteins with CC agronomic, industrial or pharmaceutical importance, including production CC of vaccines, healthcare products like human haemoglobin, industrial or CC household enzymes. Plants which can be transformed with the constructs CC of the invention include maize, millet, sorghum, sugar cane, rice, CC wheat, barley, oat, rye or turf grass.

XX

SQ Sequence 227 BP; 58 A; 42 C; 55 G; 72 T; 0 other;

Query Match 51.4%; Score 93.6; DB 21; Length 227;

Best Local Similarity 75.0%; Pred. No. 3e-21;

Matches 117; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 1 GAGCTGCTCCGGCGCTGGTCAATGAGAATGATGGAGCTGGGATTGACGTG 60

Db 1 GAGCTGCTCCGGCGCTGGTCAATGAGAATGATGGAGCTGGGATTGACGTG 60

Qy 61 AGGGGCCAGGGATGGCTATTTCTGGGAGGGACCAACGGTTCCCACTAGAAATA 120

Db 61 AGGGGCCAGGGATGGCTATTTCTGGGAGAACTAACCGATGTCAGGGACATT 120

Qy 121 ATTTGTTAACTTAACGGGATATACTATGG 156

Db 121 TATTTAACATTGATAATTGCAAAACATTTC 156

Query Match 51.4%; Score 93.6; DB 21; Length 227;

Best Local Similarity 75.0%; Pred. No. 3e-21;

Matches 117; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 1 GAGCTGCTCCGGCGCTGGTCAATGAGAATGATGGAGCTGGGATTGACGTG 60

Db 1 GAGCTGCTCCGGCGCTGGTCAATGAGAATGATGGAGCTGGGATTGACGTG 60

Qy 61 AGGGGCCAGGGATGGCTATTTCTGGGAGAACTAACCGATGTCAGGGACATT 120

Db 61 AGGGGCCAGGGATGGCTATTTCTGGGAGAACTAACCGATGTCAGGGACATT 120

Qy 121 ATTTGTTAACTTAACGGGATATACTATGG 156

Db 121 TATTTAAATTCGATAATTTCGAAAAACATTTCG 156
 AAZ61382 standard; DNA; 1985 BP.

RESULT 12
 AAF81268
 ID AAF81268 standard; DNA; 1049 BP.
 XX
 AC AAF81268;
 XX
 DT 05-JUN-2001 (first entry)
 XX
 DE DNA fragment designed for CRE-induced expression of recombinant protein.
 XX
 KW CRE recombinase; plastid genome manipulation;
 KW site-specific recombination; ds.
 XX
 OS Synthetic.
 XX
 PN WO200121768-A1.
 XX
 PD 29-MAR-2001.
 XX
 PP 21-SEP-2000; 2000WO-US25930.
 XX
 PR 21-SEP-1999; 99US-0155007.
 PR 13-JUN-2000; 2000US-021139.
 XX
 PA (RUTG) UNIV RUTGERS STATE NEW JERSEY.
 XX
 PI Maligna P, Cornuelle S, Lutz K;
 XX
 DR WPI; 2001-266071/27.
 XX
 PT Removing target nucleic acid sequences e.g. selectable marker genes, genes involved in plant cell metabolism, growth development and fertility from plastid genomes, by Cre-mediated site specific recombination.
 PT
 XX
 PT Example 3: Page 48; 83PP; English.
 XX
 CC The present sequence was used in an example illustrating an invention relating to a method for manipulating the genome of higher plants. The method involves selecting plant cells expressing proteins encoded by a CC construct having a nucleic acid encoding a marker, excision sites CC and plastid targeting sequence for homologous recombination into a CC plastid genome at a target sequence. The method is useful for removing CC heterologous sequences from the plastid genome, such as selectable CC marker genes following successful isolation of transformed plant CC and for removing endogenous genes associated with male sterility, ctpP CC ribosomal proteins and ribosomal RNA operon sequences from the plastid CC genome.
 CC
 XX Sequence 1049 BP; 269 A; 235 C; 282 G; 263 T; 0 other;
 SQ Query Match 51.4%; Score 93.6%; DB 22; Length 1049;
 Best Local Similarity 75.0%; Prod. No. 5.5e-22;
 Matches 117; Conservative. 0; Mismatches 39; Indels 0; Gaps 0;
 QY 1 GACCTCGTCCCCCGCCGCGTCAATGAGAATGGATAAGAGGCTCGTGGGATGACCTG 60
 1 GAGGTGCGTCGCCCCGGCTCAAGGATGGATAAGGGCTCGTGGGATGACCTG 60
 QY 61 AGGGGGCAAGGATGGCTATTTCGGAGGATATACATATGG 120
 Db 61 AGGGGGCAAGGATGGCTATTTCGGAGGATATACATATGG 156
 QY 121 ATTGTGTTAAGGAGGATATACATATGG 156
 Db 121 TATTTAAATTCGATAATTTCGAAAAACATTTCG 156
 Query Match 51.4%; Score 93.6%; DB 21; Length 1985;
 Best Local Similarity 75.0%; Prod. No. 7e-21;

ID XX AAZ61382 standard; DNA; 1985 BP.
 AC XX
 DE XX Nucleotide sequence of DNA construct FLARE16-S2.
 KW XX Green fluorescent protein; GFP; aada; aptaB; protein expression; vaccine; haemoglobin; enzyme; psDA; ss.
 OS Synthetic.
 OS Unidentified.
 OS Aequorea victoria.
 XX
 Key Location/Qualifiers
 PH misc_feature 7..221
 FT /*tag- a
 FT /*note- "atpB downstream box"
 FT misc_feature 228..1010
 FT /*tag- b
 FT /*note- "aada sequence"
 FT misc_feature 1059..1779
 FT /*tag- C
 FT /*note- "green fluorescent protein region"
 FT misc_feature 1790..1978
 FT /*tag- d
 FT /*note- "psBA region"
 PN WO200007431-A1.
 XX
 PD 17-FEB-2000.
 XX
 PR 03-AUG-1999; 99WO-US17806.
 XX
 PR 03-AUG-1998; 98US-0095163.
 PR 03-AUG-1998; 98US-0095167.
 PR 15-DEC-1998; 98US-012257.
 PR 29-APR-1999; 99US-0131611.
 PR 11-JUN-1999; 99US-0138764.
 PA (RUTG) UNIV RUTGERS STATE NEW JERSEY.
 XX
 PI Maligna P, Kuroda H, Khan MS;
 XX
 DR WPI; 2000-20525/18.
 XX
 PT New recombinant DNA constructs for expressing high levels of heterologous protein in plastids of higher plants, includes promoter, a leader sequence and a downstream box element -
 PT
 PT Disclosure; Fig 30; 164PP; English.
 XX
 CC The present sequence represents a DNA construct of the invention.
 CC The specification describes recombinant DNA constructs for expressing heterologous proteins in the plastids of higher plants. The DNA CC constructs comprise a 5' regulatory region which includes a promoter CC element, a leader sequence and a downstream box element operably linked CC to a coding region of the heterologous protein. The chimeric regulatory CC region enhances translational efficiency of an mRNA molecule encoded by CC the DNA construct. The DNA constructs are used for producing transformed CC monocot and dicot plants having high levels of heterologous protein CC expression. They can be used to drive expression of proteins with CC agronomic, industrial or pharmaceutical importance, including production CC of vaccines, healthcare products like human haemoglobin, industrial or CC household enzymes. Plants which can be transformed with the constructs CC of the invention include maize, millet, sorghum, sugar cane, rice, CC wheat, barley, oat, rye or turf grass.
 XX sequence 1985 BP; 553 A; 431 C; 482 G; 519 T; 0 other;
 SQ Query Match 51.4%; Score 93.6%; DB 21; Length 1985;
 Best Local Similarity 75.0%; Prod. No. 7e-21;

heterologous proteins in the plastids of higher plants. The DNA constructs comprise a 5' regulatory region which includes a promoter element, a leader sequence and a downstream box element operably linked to a coding region of the heterologous protein. The chimeric regulatory region enhances translational efficiency of an mRNA molecule encoded by the DNA construct. The DNA constructs are used for producing transformed monocot and dicot plants having high levels of heterologous protein expression. They can be used to drive expression of proteins with agronomic, industrial or pharmaceutical importance, including production of vaccines, healthcare products like human haemoglobin, industrial or household enzymes. Plants which can be transformed with the constructs of the invention include maize, millet, sorghum, sugar cane, rice, wheat, barley, oat, rye or turf grass.

Sequence 185 BP; 47 A; 35 C; 51 G; 52 T; 0 other;

Query Match 50.7%; Score 92.2; DB 21; Length 185;
 Best Local Similarity 88.5%; Pred. No. 8.2e-21;
 Matches 100; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Qy 1 GAGCTCGCTCCCGCGCTCGTGTGTTCAATGAAATGGATAAAGAGGCTCGTGATTGACCTG 60
 Db 1 GAGCTCGCTCCCGCGCTCGTGTGTTCAATGAAATGGATAAAGAGGCTCGTGATTGACCTG 60
 Qy 61 AGGGCCAGGATGGCTATTTCTGGAGGGACCAACGTTTCCACT 113
 Db 61 AGGGCCAGGATGGCTATTTCTGGACCAAAAGCCTCCATTTCATT 113

Search completed: May 25, 2003, 13:45:18
 Job time : 213 secs